凸分析导论

〔荷〕J.V.蒂尔 著

王 琦 译

李泽民 校



中南工业大学出版社



封面设计：郭　凤

ISBN 7-81020-282-0/O·045

定　价：1.20元



凸分析导论

[荷] J.V 蒂尔 著
王 琦 译
李泽民 校

中南工业大学出版社

内容简介

这是一本关于凸集，凸函数和凸最优化的入门书。它强调这个数学领域的基本概念与独特方法。定理证明着重突出凸分析的基本思想。每章有大量练习（在书末有答案和提示）以帮助读者理解所学的概念，并作简要的探索。

本书供数学、物理、工程、控制论和经济等领域高年级大学生及研究生作为教材，对于需要凸分析导引的科技工作者也是一本可取的参考书。

凸分析导论

[荷] J. V. 蒂尔 著

王 靖译

李泽民校

责 任 编 辑：程葆藤

插图责任编辑：刘楷英

*

中南工业大学出版社 出版发行

中南工业大学出版社印刷厂印装

湖南省新华书店经销

*

开本：787×1092/32 印张：6 字数：140千字

1990年1月第1版 1990年1月第1次印刷

印数：0001—1500

*

ISBN 7-81020-282-0/O·045

定价：1.20元

序

本书是我在荷兰乌得勒支大学讲授凸分析的基础上写成的。目前人们在理论及应用方面对凸分析越来越感兴趣。考虑到本书是一本入门教材，因此，我力求着重强调凸分析的基本概念和这一数学分支所特有的方法（诸如分离，次梯度，共轭函数，凸最优化等）。各章末尾的大量基本练习（在书末有答案和提示）是帮助读者理解所使用的这些概念的。

本书可供对凸性理论有兴趣的青年朋友使用，他们已有微积分，线性代数和一般拓扑学的数学基础；并熟悉泛函分析的基本概念（例如赋范线性空间，希尔伯特空间，对偶等）。

为了表达本书的特色和引起学生的兴趣，我不限于在实践中经常要处理的有穷维情形。但为使局部凸空间—凸分析的本质部份—的主题尽可能地简单，本书仅限于赋范空间。

在书目摘要注释中，我们收集了一些历史性的评论和其他材料，当然，这些决不是详尽的。

第一章我们概括了实直线上的实凸函数理论的要点；对取无穷值的函数也作了一些推广。

第二章研究线性空间中凸集的代数性质。对于线性拓扑空间，我们给出了凸集的某些拓扑特性。

第三章把线性空间中的分离定理拓广到线性拓扑空间，得出 Hahn-Banach 定理。

第四章介绍 R^n 中有关凸子集的一些经典定理及对多面锥的某些应用。我们利用相对内部的概念研究了 R^n 中的分离。

第五章研究定义在线性空间上可以取无穷值的函数。在某种意义上，局部有界性原来等价于连续性。还研究了重要的下半连续性和次可微性的概念。

第六章叙述对偶理论，并给出二次极函数和支撑函数的某些特性。

第七章给出在最优化中，凸性的内涵的一些刻划。本章主要讨论凸规划（Kuhn-Tucker 条件，鞍点和 Fenchel 的对偶定理）。

我衷心感谢 John Horváth 教授，他建议把我讲课的笔记写成英译本。我的同事 Tineke de Bunje 和 Leen Roozemond 曾读过我的全部或部分手稿，并作了许多修改，在此，特向他们致谢。最后，对 M.M.Meije 夫人花费许多时间打印手稿表示感谢。

Jan Van Tiel

目　　录

第一章　实直线上的凸函数

这一章里，我们用 I 表示 R 中的（闭，开或半开，有限或无限）区间。

实凸函数

1.1　定义

设函数 f：$I \to \mathrm{R}$。

(a) f 称为是凸的，如果对任何的 a，$b \in I$ 及任意实数 $\lambda \in \langle 0, 1 \rangle$，有

$$f(\lambda a+(1-\lambda) b) \leqslant \lambda f(a)+(1-\lambda) f(b). \qquad (1)$$

图 1 表明凸性的几何意义：端点为 $(a, f(a))$ 和 $(b, f(b))$ 的弦总是位于 f 的图形的上方。

(b) f 称为是严格凸的，如果 f 为凸的，且在（1）式中，对无论怎样的 $a \neq b$ 均有严格不等式成立。

1.2

我们给出 f：$I \to \mathrm{R}$ 呈凸性的几个其他等价描述：

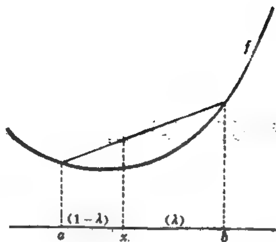


图 1

(a) 对任意 $a, b, x\in I$, $a<x<b$, 有

$$f(x)\leqslant \frac{b-x}{b-a}f(a)+\frac{x-a}{b-a}f(b)$$

注意这个不等式的右边可以写成

$$f(a)+\frac{f(b)-f(a)}{b-a}(x-a)$$

(b) 对任何的 $a, b\in I$ 及任意 $\lambda, \mu\in \mathbf{R}$, 且 $\lambda>0, \mu>0$, $\lambda+\mu=1$, 有

$$f(\lambda a+\mu b)\leqslant \lambda f(a)+\mu f(b)$$

1.3

以下简单性质的证明留给读者。

(a) 如果 f 和 g 是凸函数, 且 $\alpha\geqslant 0$, $\beta\geqslant 0$, 则 $\alpha f+\beta g$ 也是凸函数。

(b) 有限多个凸函数的和是凸函数。

(c) 收敛的凸函数序列的（点态）极限是凸函数。

(d) 设 $f: I \to \mathbf{R}$ 是凸的。则有

$$\sum_{i=1}^{n} \lambda_i x_i \in I \qquad f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leqslant \sum_{i=1}^{n} \lambda_i f(x_i)$$

其中 $x_i \in I$，$\lambda_i \geqslant 0$ $(1 \leqslant i \leqslant n)$，$\sum_{i=1}^{n} \lambda_i = 1$。

(e) 设 f 是任意多个凸函数 $I \to \mathbf{R}$ 的点态上确界，若 f 在 I 上处处有限，那么 f 是凸的。对于下确界，类似的命题成立吗？

1.4 定理

设 $f: I \to \mathbf{R}$ 是凸的，则对任何的 $a, b, x \in I$，$a < x < b$，有

$$\frac{f(x)-f(a)}{x-a} \leqslant \frac{f(b)-f(a)}{b-a} \leqslant \frac{f(b)-f(x)}{b-x} \qquad (2)$$

如果 f 是严格凸的，则（2）式中的严格不等式成立。

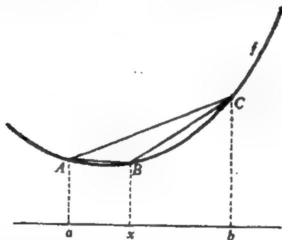


图 2

图 2 表明这个定理的几何意义：

斜率(AB)≤斜率(AC)≤斜率(BC)。

证明。因为 f 是凸的，我们有

$$f(x)\leqslant\frac{b-x}{b-a}f(a)+\frac{x-a}{b-a}f(b) \tag{3}$$

从这个不等式我们可以推出

$$f(x)-f(a)\leqslant\frac{a-x}{b-a}f(a)+\frac{x-a}{b-a}f(b)$$

$$=\frac{x-a}{b-a}[f(b)-f(a)]$$

这证明了（2）式中的第一个不等式；第二个不等式可用类似方法证明。如果 f 是严格凸的，则在（3）式中严格不等式成立，在（2）式中亦如此。

1.5

我们用 $\mathrm{int}(I)$ 表示 I 的内部。设 $f: I\to\mathbf{R}$ 是凸的，$c\in\mathrm{int}(I)$，假定 $[a, b]\subset I$，满足 $a<c<b$。由定理1.4，对任意 $x\in\langle c, b]$，我们有

$$\frac{f(c)-f(a)}{c-a}\leqslant\frac{f(x)-f(c)}{x-c}$$

由定理1.4也得出函数

$$x\to\frac{f(x)-f(c)}{x-c}$$

在 $\langle c, b]$ 上是非减的。因此，右导数

$$f'_+(c):=\lim_{x\downarrow c}\frac{f(x)-f(c)}{x-c}$$

存在。用类似方法可以证明左导数 $f'_-(c)$ 存在。

如果 $a<c<d<b$，则对于充分小的正数 h，我们有

$$\frac{f(c)-f(c-h)}{h} \leqslant \frac{f(c+h)-f(c)}{h} \leqslant \frac{f(d)-f(d-h)}{h}$$

令 $h \to 0$，取极限，我们得

$$f'_-(c) \leqslant f'_+(c) \leqslant f'_-(d).$$

这样，我们证明了下面定理.

1.6 定理

设 $f: I \to \mathbb{R}$ 是凸的，则 f 在 $\mathrm{int}(I)$ 的每一点有右导数和左导数，且 f'_- 和 f'_+ 在 $\mathrm{int}(I)$ 上是非减的. 如果 $c \in \mathrm{int}(I)$，我们有

$$f'_-(c) \leqslant f'_+(c)$$

并且，对所有的 $x \in I$（见图 3）有

$$f(x) \geqslant f(c) + f'_-(c)(x-c), \quad f(x) \geqslant f(c) + f'_+(c)(x-c).$$

注. 设 $f: [a, b] \to \mathbb{R}$ 是凸的. 上述证明指出，在这种情况，若 $+\infty$ 和 $-\infty$ 允许作为极限，那么 $f'_+(a)$ 和 $f'_-(b)$ 存在.

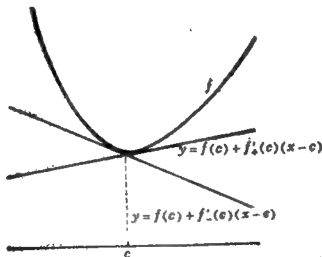


图 3

1.7

$f: I\to \mathbf{R}$ 称为是相对于 $I_0\subset I$ 的李普西兹函数，如果存在 $K>0$，使得对所有 $x, y\in I_0$，有

$$|f(x)-f(y)|\leqslant K|x-y|.$$

此条件蕴涵着 f 相对于 I_0 是连续的，甚至是一致连续的，且在 I_0 的每一个有界闭子区间上，f 是有界变差函数。

定理. 设 $f: I\to\mathbf{R}$ 是凸的，$[a, b]\subset \mathrm{int}(I)$。则

(a) f 是相对于 $[a, b]$ 的李普西兹函数。

(b) f 在 $\mathrm{int}(I)$ 连续。

证明. 显然存在 $c, d\in I$，使得 $c<a<b<d$。由定理1.6，我们有

$$f'_+(a)\leqslant f'_+(x)\leqslant \frac{f(x)-f(y)}{x-y}\leqslant f'_-(y)\leqslant f'_-(b),$$

其中 $a\leqslant x<y\leqslant b$。由此得出 $|f(x)-f(y)|\leqslant K|x-y|$，此处 $K:=\max(|f'_+(a)|, |f'_-(b)|)$。(a)得证；(b)是(a)的直接结果。

注. 即使 f 有界，f 亦未必是相对于 I 的李普西兹函数；即使 I 是闭的和有限的，f 也未必在 I 上连续。

1.8

一个相对于区间 $[a, b]$ 的李普西兹函数在 $[a, b]$ 上是绝对连续的，众所周知，这样的函数几乎处处可微，于是，从§1.7 得出凸函数几乎处处可微。

下面我们在不利用绝对连续概念的情形下，去证明凸函数

的一个较强的可微性性质。

定理 设 $f: I\to \mathbf{R}$ 是凸的，则

(a) 在 $\operatorname{int}(I)$，f'_- 左连续；f'_+ 右连续。

(b) f 仅有可数多个不可微点。

证明 (a) 由于 f 在 $\operatorname{int}(I)$ 的连续性（§1.7），对所有 x，y，$z\in \operatorname{int}(I)$，我们有

$$\frac{f(y)-f(x)}{y-x}=\lim_{z\downarrow x}\frac{f(y)-f(z)}{y-z}\geqslant\lim_{z\downarrow x}f'_+(z)$$

其中 $x<z<y$。取极限 $y\downarrow x$，得到

$$f'_+(x)\geqslant\lim_{z\downarrow x}f'_+(z)。$$

另由 f'_+ 是非减的（定理1.6），又有

$$f'_+(x)\leqslant\lim_{z\downarrow x}f'_+(z)。$$

从而 $f'_+(x)=\lim\limits_{z\downarrow x}f'_+(x)$，即 f'_+ 右连续。f'_- 的左连续性可类似证明。

(b) 由定理1.6，对所有的 x，y，$z\in \operatorname{int}(I)$，且 $x<y<z$ 有

$$f'_+(x)\leqslant f'_-(y)\leqslant f'_+(z)。$$

如果 f'_+ 在 y 连续，就有

$$f'_+(y)=\lim_{x\uparrow y}f'_+(x)=\lim_{z\downarrow y}f'_+(z)=f'_-(y)$$

这意味着 f 在 y 是可微的。从而推出 f 在 $\operatorname{int}(I)$ 的不可微点是非减函数 f'_+ 的跳跃点，由于 f'_+ 仅有可数多个这样的点，(b) 得证。

中 点 凸 性

1.9

以下概念紧密地联系到凸性

定义。函数 $f: I\to \mathrm{R}$ 称为是中点凸的，如果对所有的 a，$b\in I$，有

$$f\left(\frac{a+b}{2}\right)\leqslant\frac{1}{2}[f(a)+f(b)]. \qquad (4)$$

图4表明中点凸性的几何意义：连接 f 图形上两点的弦之中点总位于图形上对应点的上方。

1.10 定理

设 $f: I\to \mathrm{R}$ 为中点凸且连续，则 f 是凸的。

证明 假定 (a_k) 是 I 中一个序列，由（4）得出

$$f\left(\frac{a_1+a_2+a_3+a_4}{4}\right)\leqslant\frac{1}{2}f\left(\frac{a_1+a_2}{2}\right)+\frac{1}{2}f\left(\frac{a_3+a_4}{2}\right)$$

$$\leqslant\frac{1}{4}[f(a_1)+f(a_2)+f(a_3)+f(a_4)]$$

用归纳法，可以证明：对所有形如 2^k 的 n 有

$$f\left(\frac{a_1+a_2+\cdots+a_n}{n}\right)\leqslant\frac{1}{n}\sum_{i=1}^{n}f(a_i). \qquad (5)$$

现在假定当 $n=N$ 时，（5）式成立。令

$$a_N=\frac{1}{N-1}\ (a_1+a_2+\cdots+a_{N-1}),$$

我们有

$$a_N=\frac{1}{N}\ (a_1+\cdots+a_N)$$

于是

$$f(a_N)=f\left(\frac{a_1+a_2+\cdots+a_N}{N}\right)\leqslant\frac{1}{N}\sum_{i=1}^{N-1}f(a_i)+\frac{1}{N}f(a_N).$$

可见

$$f(a_N)\leqslant\frac{1}{N-1}\sum_{i=1}^{N-1}f(a_i)$$

因此，当 $n=N-1$ 时（5）式也成立. 从而导得对所有的 $n\in\mathbf{N}$（5）式成立。

设 a，$b\in I$ 及 k，$n\in\mathbf{N}$，$k<n$，从（5）式有

$$f\left(\frac{k}{n}a+\frac{n-k}{n}b\right)\leqslant\frac{1}{n}[kf(a)+(n-k)f(b)].$$

因而，对任何的分数 $\lambda\in\mathbf{Q}$，且 $\lambda\in(0,1)$

$$f(\lambda a+(1-\lambda)b)\leqslant\lambda f(a)+(1-\lambda)f(b)\qquad(6)$$

成立。鉴于 f 的连续性我们便得出对任意实数 $\lambda\in\langle 0,1\rangle$，（6）式也成立。

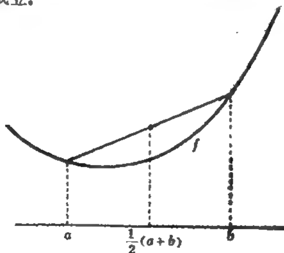


图 4

可微凸函数

1.11 定理

设 I 是开的，$f: I\to\mathbb{R}$ 二次可微，则 f 是凸的，当且仅当对所有 $x\in I$，$f''(x)\geqslant 0$。

证明 必要性。由定理1.6，f' 在 I 上是非减的，因此，对所有 $x\in I$，$f''\geqslant 0$。

充分性。设 x，$y\in I$，$x<y$，$0<\lambda<1$。按照微分学的中值定理，存在 ξ_1，ξ_2，$x<\xi_1<\lambda x+(1-\lambda)y<\xi_2<y$，和 ξ_3，$\xi_1<\xi_3<\xi_2$，使得

$$\begin{aligned}
&f(\lambda x+(1-\lambda)y)-\lambda f(x)-(1-\lambda)f(y)\\
&\quad=\lambda[f(\lambda x+(1-\lambda)y)-f(x)]+(1-\lambda)[f(\lambda x\\
&\qquad+(1-\lambda)y)-f(y)]=\lambda(1-\lambda)(y-x)f'(\xi_1)\\
&\qquad+(1-\lambda)\lambda(x-y)f'(\xi_2)\\
&\quad=\lambda(1-\lambda)(y-x)(\xi_1-\xi_2)f''(\xi_3)\leqslant 0.
\end{aligned}$$

由此得出 f 是凸的。

注。 从上述证明，我们可看出：若对所有 $x\in I$ 有 $f''(x)>0$，则 f 是严格凸的。但其逆不真，例如，函数 $f: x\to x^4$ 在 $\mathbb{R}$ 上是严格凸的，然而有 $f''(0)=0$。

1.12 不等式

从定理1.11可以得到凸函数的许多简单例子。借助于这些函数，我们能导出一些乍看起来并非很简单的不等式。我们给出一个例子：对任意的 $x>0$，$y>0$，$\lambda>0$，$\mu>0$ 及 $\lambda+\mu=1$ 有

$$x^{\lambda}y^{\mu}\leqslant\lambda x+\mu y \tag{7}$$

利用函数 $x\to e^x$ 的严格凸性，可得

$$\exp(\lambda\log x+\mu\log y)\leqslant\lambda\exp(\log x)+\mu\exp(\log y).$$

（7）的常见形式尚有

$$x^{1/p}y^{1/q}\leqslant\frac{1}{p}x+\frac{1}{q}y \tag{8}$$

及

$$xy\leqslant\frac{1}{p}x^p+\frac{1}{q}y^q \tag{9}$$

其中 $x>0$，$y>0$，$p>1$，$q>1$ 且 $1/p+1/q=1$。当 $p=q=2$ 时，（8）是熟知的不等式

$$\sqrt{xy}\leqslant\frac{1}{2}(x+y).$$

与积分有关的定理

1.13 定理

设函数 $f:\langle a,b\rangle\to\mathbf{R}$。$f$ 是凸的当且仅当 f 可表示为

$$f(x)=f(c)+\int_c^x g(t)dt\qquad(c,\ x\in\langle a,b\rangle) \tag{10}$$

其中 g 是定义在 $\langle a,b\rangle$ 上的非减右连续实值函数。

证明。必要性。设 f 是凸的，c，$x\in\langle a,b\rangle$。由定理1.6和§1.8，f'_+ 存在且为非减的右连续函数。令

$$h(\varepsilon):=\int_c^x\frac{f(t+\varepsilon)-f(t)}{\varepsilon}dt.$$

我们有

$$\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}[f(t+\varepsilon)-f(t)]=f'_+(t) \qquad (a<t<b).$$

由§1.7，存在$K>0$，使得对所有的$t\subset\langle c, x\rangle$，及所有充分小的$\varepsilon>0$，有

$$\left|\frac{1}{\varepsilon}[f(t+\varepsilon)-f(t)]\right|\leqslant K.$$

应用勒贝格控制收敛定理得到

$$\lim_{\varepsilon\downarrow 0}h(\varepsilon)=\int_c^x f'_+(t)dt$$

（注意鉴于被积函数的单调性，后面这个积分是黎曼可积的）。我们又有

$$\frac{1}{\varepsilon}\int_c^x[f(t+\varepsilon)-f(t)]dt=\frac{1}{\varepsilon}\left[\int_{c+\varepsilon}^{x+\varepsilon}f(t)dt-\int_c^x f(t)dt\right]$$

$$=\frac{1}{\varepsilon}\int_x^{x+\varepsilon}f(t)dt-\frac{1}{\varepsilon}\int_c^{c+\varepsilon}f(t)dt$$

$$\to f(x)-f(c) \qquad 当\ \varepsilon\downarrow 0$$

（由于f的连续性）。因此

$$f(x)-f(c)=\int_c^x f'_+(t)dt \tag{11}$$

充分性。假定（10）成立，其中g是非减的。设$x, y\in\langle a, b\rangle$，$x<y$及$\lambda\in\langle 0, 1\rangle$，令$z_1=\lambda x+(1-\lambda)y$，我们有（见图5）

$$\begin{aligned}&f(z)-\lambda f(x)-(1-\lambda)f(y)\\&=\lambda[f(z)-f(x)]-(1-\lambda)[f(y)-f(z)]\\&=\lambda\int_x^z g(t)dt-(1-\lambda)\int_z^y g(t)dt\\&\leqslant\lambda(z-x)g(z)-(1-\lambda)(y-z)g(z)=0\end{aligned}$$

(因为$z-x=(\lambda-1)(x-y)$及$y-z=\lambda(y-x)$)。

注。 我们知道一个函数f：$[a,b]\to\mathrm{R}$是绝对连续的，当且仅当f能表示成

$$f(x)=c+\int_a^x g(t)dt$$

其中$c\in\mathrm{R}$，g是勒贝格可积的，于是我们几乎处处有$f'(x)=g(x)$，因此(11)式是f的绝对连续性的一个直接结果（见§1.8）。

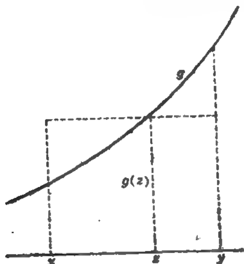


图5

1.14

设f：$[a,b]\to\mathrm{R}$是凸的，$a_i\in[a,b]$（$1\leqslant i\leqslant n$），则我们有

$$f\left(\frac{1}{n}\sum_{i=1}^{n}a_i\right)\leqslant\frac{1}{n}\sum_{i=1}^{n}f(a_i). \tag{12}$$

(12)式是n个数的算术平均值(a.m)定理：

$f(\text{a.m.of}\ a_1,a_2,\cdots,a_n)\leqslant\text{a.m.of}\ f(a_1),f(a_2),\cdots,f(a_n)$。

对于函数的平均值也有类似定理：

定理（Jensen 不等式）．设 f：$\langle a, b\rangle \to \mathbf{R}$ 是凸的，g：$[c, d] \to \langle a, b\rangle$ 是连续的．则

$$f\left(\frac{1}{d-c}\int_c^d g(x)\,dx\right) \leqslant \frac{1}{d-c}\int_c^d f(g(x))\,dx.$$

证明．令

$$p := \frac{1}{d-c}\int_c^d g(x)\,dx$$

我们有 $p \in \langle a, b\rangle$。由定理1.6，对任意的 $y \in \langle a, b\rangle$，有

$$f(y) \geqslant f(p) + f'_+(p)(y-p)$$

因此，对任意的 $x \in [c, d]$

$$f(g(x)) \geqslant f(p) + f'_+(p)[g(x) - p]$$

在 $[c, d]$ 上积分上述不等式，即得要证明的结果。

注。

(a) 在此定理中，可以用任意一个在 $[c, d]$ 上勒贝格可积的函数来代替 g。

(b) Jensen 不等式在概率论中有下述类似结论，它可用同法证明．

设 X 是一个具有概率测度 μ 的概率空间（因此 $\mu(X) = 1$）．f：$\langle a, b\rangle \to \mathbf{R}$ 是凸的，g：$X \to \langle a, b\rangle$ 是 μ—可积．则

$$f\left(\int_X g\,d\mu\right) \geqslant \int_X (f\circ g)\,d\mu.$$

用概率术语：如果 x 是 X 上的随机变量，那么我们有 $f(Ex) \leqslant E[f(x)]$，其中 Ex 是 x 的期望值。

共 轭 函 数

1.15 定理

函数 f：$\mathbf{R} \to \mathbf{R}$ 是凸的，当且仅当存在一个函数 g：$\mathbf{R} \to \mathbf{R} \cup \{+\infty\}$，使得对所有 $x \in \mathbf{R}$ 有

$$f(x)=\sup_{y\in R}[xy-g(y)]$$

证明。充分性。我们有

$$f(x)=\sup_{g(y)<+\infty}[xy-g(y)]$$

我们看出 f 是一仿射函数（因此是凸的）族的点态上确界，由 §1.3(e)，f 是凸的。

必要性。定义函数 g：$R\to R\cup\{+\infty\}$为

$$g(y)=\sup_{x\in R}[xy-f(x)].$$

设 $x_0\in R$，对任意的 $y\in R$ 有

$$g(y)\geqslant x_0y-f(x_0)$$

因此

$$x_0y-g(y)\leqslant f(x_0).$$

从而

$$\sup_{y\in R}[x_0y-g(y)]\leqslant f(x_0).\tag{13}$$

令 $y_0:=f'_+(x_0)$。按照定理1.6，对任意 $x\in R$ 有

$$f(x)\geqslant f(x_0)+f'_+(x_0)(x-x_0)=f(x_0)+y_0(x-x_0)$$

因此

$$xy_0-f(x)\leqslant x_0y_0-f(x_0).$$

于是

$$g(y_0)=x_0y_0-f(x_0)$$

即

$$x_0y_0-g(y_0)=f(x_0).\tag{14}$$

由（13）和（14）式证得必要性。

1.16

上面的 g 称为 f 的共轭函数。f 和 g 组成函数对，使对所

有 x，$y\in\mathbb{R}$ 满足不等式

$$f(x)+g(y)\geqslant xy. \tag{15}$$

我们给出定理1.15的下述几何解释（见图6）。斜率为 y，截距为 $-a$ 的直线 m 处处位于 f 的图形之下方当且仅当对任意的 $x\in\mathbb{R}$，有

$$yx-a\leqslant f(x)$$

即是

$$a\geqslant xy-f(x).$$

满足这个不等式的最小数 a 是

$$\sup_{x\in\mathbb{R}}[xy-f(x)]=g(y)$$

因此，尽可能地向上平移 m，以得到一条与 f 的图形相交，且截距等于 $-g(y)$ 的直线 $n(y)$。定理1.15告诉我们 f 的图形是这类直线 $n(y)$ $(y\in\mathbb{R})$ 的包络当且仅当 f 是凸的。建议读者去给出定理证明中选得的 y_0 的几何解释，并给出'$g(y)=+\infty$'的几何意义。

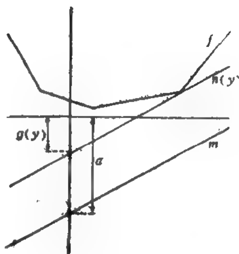


图6

1.17 例

(a) 设 $p>1$，$f(x)=|x|^p/p$ $(x\in\mathbb{R})$，则

$$g(y)=\frac{1}{q}|y|^q$$

此处$\frac{1}{p}+\frac{1}{q}=1$。因此，由（15）式，对所有实数$x$和$y$（见§1.12）有

$$xy\leqslant \frac{1}{p}|x|^p+\frac{1}{q}|y|^q. \tag{16}$$

(b) 设f：$[0,\infty>\to\mathbf{R}$为严格增的连续函数，且$f(0)=0$。设g是f的反函数，我们定义F和G为

$$F(x)=\begin{cases}\int_0^x f(t)dt & \text{如果 } x\geqslant 0\\ 0 & \text{如果 } x<0\end{cases}$$

$$G(y)=\int_0^y g(u)du \qquad \text{如果 } y\in f([0,\infty>).$$

由定理1.13，F和G为凸函数；可以证明G是F的共轭在$f([0,\infty>)$上的限制。因此对每个$x\geqslant 0$，$y\in f([0,+\infty>)$，有

$$xy\leqslant F(x)+G(y)$$

因此

$$xy\leqslant \int_0^x f(t)dt+\int_0^y g(u)du. \tag{17}$$

（17）叫做Young不等式。图7给出它的几何解释，并指出初等证明的思路。在Young不等式中，取$f(x)=x^{p-1}$，$g(y)=y^{q-1}$（这里$p>1$且$\frac{1}{p}+\frac{1}{q}=1$），我们再一次得（16）式。

在R上取值的凸函数

1.18

在定理1.15中已涉及到在$\mathbf{R}\cup\{+\infty\}$中取值的凸函数。

今后我们将考虑更一般的取值于$\bar{\mathbf{R}}:=\mathbf{R}\cup\{+\infty\}\cup\{-\infty\}$中的函数。至于含有$+\infty$和$-\infty$的计算，我们采用显然的规则，例如：$x+(+\infty)=+\infty$，若$x\in\mathbf{R}$；$x\cdot(-\infty)=-\infty$，若$x>0$。也采用下述不太明显的规则：

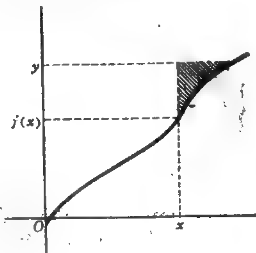


图7

$$0\cdot(+\infty)=(+\infty)\cdot 0=0\cdot(-\infty)=(-\infty)\cdot 0=0.$$

而$+\infty-\infty$未作定义。

下面，我们继续推广凸函数的概念。

1.19 定义

函数$f:\mathbf{R}\to\bar{\mathbf{R}}$称为是凸的，如果对所有满足$f(x)<\mu$，$f(y)<\nu$和$0<\lambda<1$的$x,y,\lambda,\mu,\nu\in\mathbf{R}$，有

$$f(\lambda x+(1-\lambda)y)<\lambda\mu+(1-\lambda)\nu. \qquad (18)$$

1.20

设$f:\mathbf{R}\to\mathbf{R}$是凸函数，$f(x)<\mu$，$f(y)<\nu$，$0<\lambda<1$，

则

$$f(\lambda x+(1-\lambda)y)\leqslant\lambda f(x)+(1-\lambda)f(y)<\lambda\mu+(1-\lambda)\nu。$$

反之，若（18）式成立，则对任意的 $\varepsilon>0$ 由于 $f(x)<f(x)+\varepsilon$，$f(y)<f(y)+\varepsilon$ 我们有

$$f(\lambda x+(1-\lambda)y)<\lambda f(x)+(1-\lambda)f(y)+\varepsilon$$

于是

$$f(\lambda x+(1-\lambda)y)\leqslant\lambda f(x)+(1-\lambda)f(y)$$

f 就是凸的。可见定义1.19确是定义1.1的推广。

1.21　**定义**

(a) 凸函数 f：$\mathbf{R}\to\overline{\mathbf{R}}$ 的有效域是集合

$$\{x\in\mathbf{R}\mid f(x)<+\infty\},$$

用 dom(f) 表示。

(b) $\mathbf{R}$ 上的正常凸函数是不恒等于 $+\infty$ 的凸函数 $\mathbf{R}\to\mathbf{R}\cup\{+\infty\}$。

(c) $\mathbf{R}$ 上的非正常凸函数是 $\mathbf{R}$ 上的一个凸函数且不是正常的。

1.22

读者易证凸函数 f：$\mathbf{R}\to\overline{\mathbf{R}}$ 的有效域是凸的（因此是一个区间）。

有效域为 I 的正常凸函数 F：$\mathbf{R}\to\overline{\mathbf{R}}$ 可以看作是由某个域为 I 的有限凸函数 f 扩充而得到的。对所有 $x\in\mathbf{R}$，其扩充方法如下：

$$F(x)=\begin{cases} f(x) & \text{如果 } x\in I \\ +\infty & \text{如果 } x\bar{\in} I. \end{cases}$$

这个扩充过程使我们能够把具有不同定义域的有限凸函数处理成定义在整个 R 上而取值于 $\overline{R}$ 的凸函数。

注意，定理1.15中的函数 g 在上述意义下就是凸的。

1.23

我们不难描绘非正常凸函数类。

定理. 设 f: R→$\overline{R}$ 非正常凸函数，则对任意 $x\in$ int (dom(f))，有 $f(x)=-\infty$。

证明。如果 $f\equiv+\infty$(即是对所有的 $x\in R, f(x)=+\infty$)，命题显然是成立的。如果 $f\not\equiv+\infty$，则存在 $a\in R$，使得 $f(a)=-\infty$（注意 $a\in$ dom(f))。设 $x\in$ int(dom(f))，$x\neq a$。从而存在 $y\in$ dom(f)及 $\lambda\in\langle 0,1\rangle$，使 $x=\lambda a+(1-\lambda)y$。由定义1.19，对每个满足 $f(y)<\alpha<+\infty$ 的 α，及每个 $\beta\in R$ 有

$$f(x)=f(\lambda a+(1-\lambda)y)<\lambda\beta+(1-\lambda)\alpha$$

(因为 $f(a)=-\infty<\beta$)。令 $\beta\to-\infty$，我们得 $f(x)=-\infty$。

1.24

在 § 1.3 中所叙实凸函数的性质，对于取值于 $\overline{R}$ 的凸函数是类似的，但对其中的 (a)，(b) 和 (d) 我们限于正常凸函数（为了避免象 $+\infty-\infty$这样的式子）.

下文将用到这个事实：只要允许取值 $+\infty$和 $-\infty$，那么，R上的正常凸函数在整个有效域上有右导数和左导数。我们就 dom (f) =[a,b]的情形给出证明.由 § 1.5，对任何 $x\in[a,b,\rangle$

$f'_{+}(x)$ 存在，对任何 $x\in\langle a, b]$，$f'_{-}(x)$ 存在。对 $x<a$，我们有 $f(x)=+\infty$，则由

$$\frac{f(x)-f(a)}{x-a}=-\infty$$

就有 $f'_{-}(x)=-\infty$；而对任何 $x>b$，我们又有

$$\frac{f(x)-f(b)}{x-b}=+\infty$$

此即　$f'_{+}(b)=+\infty$。

推　　广

1.25　拟凸性

某些概念与凸性紧密相关。我们举出一个例子

定义。 设函数 $f: I\to\mathbf{R}$。

(a) f 称为是拟凸的，如果对任何的 $a, b\in I$，$f(a)\leqslant f(b)$ 及所有的 $\lambda\in\langle 0, 1\rangle$，有

$$f(\lambda a+(1-\lambda)b)\leqslant f(b)。$$

(b) f 称为是严格拟凸的，如果对任何的 $a, b\in I$，$f(a)<f(b)$ 及所有的 $\lambda\in\langle 0, 1\rangle$，有

$$f(\lambda a+(1-\lambda)b)<f(b)。$$

严格拟凸函数并非一定是拟凸的（见练习10）。

练　　习

1　设函数 $f: \langle a, b\rangle\to\mathbf{R}$。证明下述命题：

(a) 若 f是凸函数，则 f是单调的（非减或非增）或存在 $c\in\langle a, b\rangle$，使得 f在 c 的左边非增，而在 c 的右边是非减的。

(b) 若 f是凸的，则 f的每个局部（相对）极小点必是整体（绝对）极小点。

(c) 若 f是严格凸的，则 f最多有一个整体极小点。

2 设 f：$\langle a, b\rangle\to\mathbf{R}$ 是凸的，且 $c\in\langle a, b\rangle$。证明：f在 c 可微当且仅当

$$\lim_{h\downarrow 0}\frac{f(c+h)+f(c-h)-2f(c)}{h}=0\text{。}$$

3 设 f：$\langle a, b\rangle\to\mathbf{R}$。证明：$f$是凸的，当且仅当对 f 图形上的每点 p，至少存在一条通过 p 且位于 f图形下方的直线。

4 (a) 设 x_1，…，x_n，r_1，…，r_n是正实数，满足 $\sum_{i=1}^{n} r_i=1$。证明

$$\prod_{i=1}^{n} x_i^{r_i}\leqslant\sum_{i=1}^{n} r_i x_i\text{。}$$

(b) 证明：n 个正实数的几何平均值不大于它们的算术平均值。

(c) 证明 Hölder 不等式：

$$\sum_{i=1}^{n} a_i b_i\leqslant\left(\sum_{i=1}^{n} a_i^p\right)^{1/p}\left(\sum_{i=1}^{n} b_i^q\right)^{1/q}$$

此中 a_1，a_2，…，a_n，b_1，b_2，…，b_n是非负实数，且 $p>1$，$q>1$，$1/p+1/q=1$。取 $p=q=2$，我们得到柯西不等式。

5 证明有界凸函数 f：$\mathbf{R}\to\mathbf{R}$ 必是常数。

6 正函数f被称为是对数凸的（$l.c.$），是指$\log f$是凸的。设f和g是二次可微函数$\mathbf{R}\to\mathbf{R}$，证明下述命题：

(a) f是对数凸的，当且仅当$f>0$且$f''f\geqslant(f')^2$。

(b) 如果f是对数凸的，则f是凸的。

(c) 如果f和g是对数凸的，且$\alpha>0$，$\beta>0$，则$\alpha f+\beta g$也是对数凸的。

(d) 设a_1，a_2，…，a_n是正实数。则函数

$$x\to\log(a_1^x+\cdots+a_n^x)$$

在$\mathbf{R}$上是凸的。

7 设f和g是凸函数$\mathbf{R}\to\mathbf{R}$，若f是非减，证明函数$x\to f(g(x))$是凸的。

8 证明：在§1.17例（b）中，函数G是F的共轭的限制。

9 设函数f：$\mathbf{R}\to\mathbf{R}\cup\{+\infty\}$。证明：$f$是凸的，当且仅当对任何$a$，$b\in\mathbf{R}$及任何$\lambda\in\langle0,1\rangle$，有

$$f(\lambda a+(1-\lambda)b)\leqslant\lambda f(a)+(1-\lambda)f(b).$$

10 设函数f：$\langle a,b\rangle\to\mathbf{R}$。

(a) 证明：f是拟凸的，当且仅当对每个$\alpha\in\mathbf{R}$，集$\{x\in\langle a,b\rangle\,|f(x)\leqslant\alpha\}$是凸的。

(b) 证明严格拟凸性不蕴涵拟凸性。

(c) 若f是连续的严格拟凸函数，证明f是拟凸的。

注　　释

1 从定理1.10得出，对于连续函数，凸性与中点凸性是等价的。此结论见

J. L. W. V. Jensen, Sur les fonctions convexes et les inégalités entre les valeurs moyennes, *Acta Math.* 30(1906) 175—93.

一些更好结果是在下列文章和这些文章所引用的著作中.

H. Blumberg, On convex functions, *Trans. Am. Math. Soc.* 20(1919) 40—44,

W. Sierpiński, Sur les fonctions convexes mesurables, *Fund. Math.* 1 (1920) 125-8,

A. Ostrowski, Zur Theorie der Konvexen Funktionen, *Comm. Math. Helvetici* 1 (1929) 157-9,

例如可测的中点凸函数是凸的（因此是连续的）便是其中之一.

2 读者可在下面书中见到更多的不等式（见§1.12）

G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, Cambridge, Cambridge University Press, 1934.

3 对所有 $x,y\in R$，满足不等式

$$f(x)+g(y)\geqslant xy$$

的函数对 f, g（见§1.16）的最早论述给出在下文中

Z. W. Birnbaum and W. Orlicz, Über die Verallgemeinerung des Begriffes der zueinander Konjugierten Potenzen, *Studia Math.* 3 (1931) 1-67.

这也是第一次使用共轭函数概念的论文（在复变函数论的文献中，“共轭”有着与此不同的含意）.关于 Birnbaum 和 Orlicz 文章中的基本结果在下面专著的第一章中也可找到

M. A. Krasnosel'skii and Ya. B. Rutickii, *Convex functions and Orlicz spaces*, Groningen, Noordhoff, 1961.

Young 在下文中就使用了不等式 (17) (见 § 1.17)

W. H. Young, On classes of summable functions and their Fourier series, *Proc. R Soc.* A 87(1912)225-9. 定理1.15被给出于文

S. Mandelbrojt, Sur les fonctions convexes, *C. R. Acad. Sci.* 209 (1939) 977-8.

4 函数 f: R→R 的凸性可以表述如下：对任何 $x<y<z$，成立

$$\begin{vmatrix} 1 & 1 & 1 \\ x & y & z \\ f(x) & f(y) & f(z) \end{vmatrix} \geqslant 0$$

其推广形式为

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_{n+1} \\ x_1^2 & x_2^2 & \cdots & x_{n+1}^2 \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_{n+1}^{n-1} \\ f(x_1) & f(x_2) & \cdots & f(x_{n+1}) \end{vmatrix} \geqslant 0$$

可在下述文献中找到

S. Karlin and Z. Ziegler, Some applications to inequalities of the method of generalized convexity, *J. d'-Analyse Math*, 30 (1976) 281-303.

第二章　线性空间中的凸子集

本章我们将用V表示R上的线性空间，为简单起见，假设V中不止含一个点。

凸包和仿射包

2.1　定义

(a)　设x，$y\in V$。线段$[x,y]$（具有端点x和y）是集$\{\lambda x+(1-\lambda)y\,|\,0\leqslant\lambda\leqslant 1\}$。如果$x\neq y$，$[x,y]$的内部$\langle x,y\rangle$是集$\{\lambda x+(1-\lambda)y\,|\,0<\lambda<1\}$。用类似方法可定义$[x,y\rangle$和$\langle x,y]$。

(b)　设$A\subset V$。A称为凸的，如果对任何的x，$y\in A$有$[x,y]\subset A$。

(c)　$x_1,x_2,\cdots,x_n\in V$的（有限）凸组合是V中形如

$$\sum_{i=1}^{n}\lambda_i x_i$$

的点，这里$\lambda_i\geqslant 0(1\leqslant i\leqslant n)$，$\sum_{i=1}^{n}\lambda_i=1$。

2.2

读者易证任意个凸集之交仍是凸集. 因此对每个 $A\subset V$，V 中所有包含 A 的凸子集之交是 V 中包含 A 的最小凸子集.

定义 设 $A\subset V$，A 的凸包 $\mathrm{co}(A)$ 是指 V 中包含 A 的最小凸子集.

2.3 **定理**

设 $A\subset V$，$\mathrm{co}(A)$ 是 A 中元素的所有（有限）凸组合组成的集.

证明. 用 B 表示 A 中元素的所有凸组合构成的集. 因为 $\mathrm{co}(A)$ 是包含 A 的凸集，所以它包含所有端点在 A 中的线段，亦就包含 A 中元素的所有凸组合（见 § 1.3 性质(d)）. 故 $B\subset \mathrm{co}(A)$. B 是凸的. 因为 B 中两个元素

$$x=\sum_{i=1}^{n}\lambda_i x_i \quad 及 \quad y=\sum_{i=1}^{m}\mu_i y_i$$

的凸组合 $\lambda x+(1-\lambda)y$ 是 A 中元素的凸组合

$$\sum_{i=1}^{n}(\lambda\lambda_i)x_i+\sum_{i=1}^{m}[(1-\lambda)\mu_i]y_i.$$

所以，B 是包含 A 的凸集，即有 $B\supset \mathrm{co}(A)$. 这样 $B=\mathrm{co}(A)$.

2.4

下列简单性质的证明留给读者.

(a) 凸集的线性组合是凸的：若 $C_1, C_2, \cdots, C_n$ 是 V 中的凸子集，$\alpha_i\in \mathbf{R}(1\leqslant i\leqslant n)$，则 $\sum_{i=1}^{n}\alpha_i C_i$ 是凸的. 这里

$$C_1+C_2=\{x_1+x_2\,|x_1\in C_1,\ x_2\in C_2\}$$
$$\alpha C=\{\alpha x\,|x\in C\}.$$

(b) 设$A\subset V$.则A是凸的，当且仅当对任何$\alpha\geqslant 0$，$\beta\geqslant 0$有$(\alpha+\beta)A=\alpha A+\beta A$.

(c) 设W是线性空间，T: $V\to W$为线性映射，若$A\subset V$是凸的，则TA是凸的．若$B\subset W$是凸的，则$T^{-1}B$是凸的．

(d) 若A，$B\subset V$是凸的，则

$$\mathrm{co}(A\cup B)=\bigcup_{0\leqslant\lambda\leqslant 1}[\lambda A+(1-\lambda)B]$$

(e) 若$A\subset V$，$x\in\mathrm{co}(A)$，则$\mathrm{co}(A\cup\{x\})=\mathrm{co}(A)$.

2.5

设$a\in V$，L是V中线性子空间，称$a+L$为V的仿射子集（或线性流形）．$a+L$的维数$\dim(a+L)$定义为$\dim(a+L)=\dim(L)$.设$A\subset V$，我们把下述命题的等价性留给读者去证明：

(a) A是仿射集．

(b) 过A中任意两点的直线仍在A中，即$x,y\in A$蕴涵$\lambda x+(1-\lambda)y\in A$，对所有$\lambda\in R$.

V中包含A的所有仿射子集的交是V的另一仿射子集．并且是V中包含A的最小仿射子集．

2.6 定义

(a) 设$A\subset V$．A的仿射包$\mathrm{aff}(A)$是指V中包含A的最小仿射子集．

(b) A的维数$\dim(A)$就是$\mathrm{aff}(A)$的维数．

(c) 点$x_1,x_2,\cdots,x_n\in V$的（有限）仿射组合是V的点，且

可表为和式

$$\sum_{i=1}^{n}\lambda_i x_i$$

其中$\lambda_i \in R (1 \leqslant i \leqslant n)$，$\sum_{i=1}^{n}\lambda_i = 1$（见定义2.1(c)）。

2.7 定理

设$A \subset V$。$\mathrm{aff}(A)$ 是A中元素的所有仿射组合构成的集。
证明。完全类似于定理2.3那样地证明。

凸 多 胞 形

2.8

设$x_1, x_2, \cdots, x_n \in V$，我们用 $\mathrm{co}(x_1, x_2, \cdots, x_n)$ 与 $\mathrm{aff}(x_1, x_2, \cdots, x_n)$ 分别表示有限集$\{x_1, x_2, \cdots, x_n\}$的凸包与仿射包。由定理2.3和定理2.7，我们有

$$\mathrm{co}(x_1, x_2, \cdots, x_n) = \left\{\sum_{i=1}^{n}\lambda_i x_i \,\middle|\, \lambda_i \geqslant 0 \ (1 \leqslant i \leqslant n), \ \sum_{i=1}^{n}\lambda_i = 1\right\}$$

$$\mathrm{aff}(x_1, x_2, \cdots, x_n) = \left\{\sum_{i=1}^{n}\lambda_i x_i \,\middle|\, \sum_{i=1}^{n}\lambda_i = 1\right\}.$$

$\mathrm{co}(x_1, x_2, \cdots, x_n)$ 称为由 $x_1, x_2, \cdots, x_n$ 生成的凸多胞形。

2.9

设$C = \mathrm{co}(x_1, x_2, \cdots, x_n)$ 是V中的凸多胞形，如果 $x_i \in \mathrm{co}$

$(x_2, x_3, \cdots, x_n)$，由§2.4性质(c) 有$C = \mathrm{co}(x_2, x_3, \cdots, x_n)$；如果 $x_2 \in \mathrm{co}(x_3, x_4, \cdots, x_n)$，那么有 $C = \mathrm{co}(x_3, x_4, \cdots, x_n)$。如果 $x_1 \in \mathrm{co}(x_2, x_3, \cdots, x_n)$，而$x_2 \in \mathrm{co}(x_1, x_3, \cdots, x_n)$。则有$C = \mathrm{co}(x_1, x_3, \cdots, x_n)$。继续用此方法，可以得知存在$\{x_1, x_2, \cdots x_n\}$的一个具有下述性质的子集$\{a_1, a_2, \cdots, a_k\}$：

(1) $C = \mathrm{co}(a_1, a_2, \cdots, a_k)$。

(2) 每一个 a_i 都不是 $a_1, a_2, \cdots a_{i-1}, a_{i+1}, \cdots a_k$ 的凸组合。$a_1, a_2, \cdots, a_k$ 称为C的顶点，C的顶点集由（1），（2）唯一确定。

2.10 定义

设$A \subset V$，$a \in A$称为A的极端点，如果 a 不是A 中任何线段的内点。

欧几里得平面中的闭园盘与欧几里得空间中的闭球的所有边界点都是极端点。凸多胞形的极端点是它的顶点。

2.11 定义

点$x_1, x_2, \cdots, x_n \in V$称为仿射独立的，如果 $\dim(\mathrm{aff}(x_1, x_2, \cdots, x_n)) = n-1$。

定理．

设$x_1, x_2, \cdots, x_n \in V$，下述命题是等价的：

(a) $x_1, x_2, \cdots, x_n$ 是仿射独立的。

(b) 对每个 j，向量组 $x_i - x_j$，$i \neq j$，是线性独立的。

(c) $\begin{cases} \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = 0 \\ \alpha_1 + \alpha_2 + \cdots + \alpha_n = 0 \end{cases}$ （其中$\alpha_1, \cdots, \alpha_n \in \mathbb{R}$）

蕴涵 $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$。

证明. 此证明由上面定义立得。

2.12

$x_1, x_2, \cdots, x_n \in V$ 称为是仿射相关的，如果它们不是仿射独立。读者易证下述命题：

(a) 若 $x_1, x_2, \cdots, x_n$ 是仿射独立的，则 $\text{aff}(x_1, x_2, \cdots, x_n)$ 中每个点 x 可唯一地写成 $x_1, x_2, \cdots, x_n$ 的仿射组合。

(b) 若 $A \subset V$，$\dim(A) = n$ 及 $k < n+1$，则 A 中 k 个点的任意仿射独立集可被扩充为 A 中 $n+1$ 个点的仿射独立集。

(c) 假定 $A \subset V$ 且 $\dim(A) < \infty$，则 $\dim(A)$ 是使 A 包含 $n+1$ 个仿射独立点的最大整数 n。

2.13 定义

V 中的一个 k-单纯形是 V 中 $k+1$ 仿射独立点的凸包。

k-单纯形的维数是 k。0-单纯形是一个点，1-单纯形是一条线段，2-单纯形是一个三角形，3-单纯形是一个四面体等等。

如果 $C = \text{co}(a_0, a_1, \cdots, a_k)$ 是 V 中一个 k-单纯形，则 $a_0, a_1, \cdots, a_k$ 必是 C 的顶点(见§2.9)，且 C 中每个点 x 可唯一地表成 $a_0, a_1, \cdots, a_k$ 的凸组合；每个 $x \in \text{aff}(C) = \text{aff}(a_0, a_1, \cdots, a_k)$ 可以唯一地表成 $a_0, a_1, \cdots, a_k$ 的仿射组合 $\sum_{i=0}^{k} \lambda_i a_i$，其中 $\lambda_0, \lambda_1, \cdots, \lambda_k$ 称为相对于 $a_0, a_1, \cdots, a_k$ 的 x 的重心坐标。

代数内部和代数闭包

2.14 定义

设$A\subset V$。

(a) A 的代数内部A^i 由A中这样的点x组成，对V中每条通过x的直线m，它与A的交包含一条线段，使x在其内部。

(b) A的代数闭包A^a 是A与集合

$$\{x\in V \mid 存在 c\in A,\ s.t.[c,x\rangle\subset A\}$$

的并集。

在不具拓扑的线性空间中，A^i 和A^a 是有用的概念。

注。

(a) 对所有$x\in A$，我们有

$$x\in A^i \Longleftrightarrow (\forall v\in V)(存在\delta>0)[x-\delta v,\ x+\delta v]\subset A$$
$$\Longleftrightarrow (\forall v\in V)(存在\delta>0)[x,\ x+\delta v]\subset A.$$

(b) 若A为单点集，则$A^a\setminus A=\phi$。

(c) 可以证明，对任意的凸集$C\subset V$，我们有$(C^i)^i=C^i$（见§2.22），但一般仅成立$(C^a)^a\supset C^a$。

2.15 定理

若$C\subset V$是凸的，则C^i 和C^a 是凸的。

证明。(a) 设$x,y\in C^i$，$\lambda\in\langle 0,1\rangle$，$z=\lambda x+(1-\lambda)y$，及$v\in V$，则存在$\delta>0$，使用$[x,x+\delta v]\subset C$和$[y,y+\delta v]\subset C$。根据$C$的凸性，我们有

$$z+\delta v=\lambda(x+\delta v)+(1-\lambda)(y+\delta v)\in C,$$

因此$[z,\ z+\delta v]\subset C$，从而$z\in C^i$。

(b) 设$x,y\in C^a$，$\lambda\in\langle 0,1\rangle$，及$z=\lambda x+(1-\lambda)y$，则存在

$c_1,c_2\in C$,使得$[c_1,x\rangle\subset C$和$[c_2,y\rangle\subset C$。定义$c=\lambda c_1+(1-\lambda)c_2$,我们将证明$[c,z\rangle\subset C$。事实上,对任何$\mu\in\langle 0,1\rangle$,我们有

$$
\begin{aligned}
\mu c+(1-\mu)z &= \mu[\lambda c_1+(1-\lambda)c_2]+(1-\mu)[\lambda x+(1-\lambda)y] \\
&= \lambda[\mu c_1+(1-\mu)x]+(1-\lambda)[\mu c_2+(1-\mu)y] \\
&\in C\text{。}
\end{aligned}
$$

从而 $z\in C^a$。

2.16 定理

设$C\subset V$是凸的,$x\in C^i$,$y\in C^a$,则$[x,y\rangle\subset C^i$。

证明。设$\lambda\in\langle 0,1\rangle$,$z=\lambda x+(1-\lambda)y$。

(a) 首先,假定$y\in C$。设$v\in V$,则存在$\delta>0$,使$[x,x+\delta v]\subset C$。我们有 $z+\lambda\delta v=\lambda(x+\delta v)+(1-\lambda)y\in C$,因此$[z,z+\lambda\delta v]\subset C$。即$z\in C^i$。

(b) 现在假定$y\in C^a\setminus C$,设$z_1\in\langle z,y\rangle$。则存在$u\in C$,使得$[u,y\rangle\subset C$,且存在$\delta>0$,使$p=x-\delta(u-y)\in C$。若x,y,p和u不共线(其他的情形留给读者证明)。用m表示通过p和z_1的直线。对于充分小的δ,m交$[u,y\rangle$于点q(见图8)。因为$p,q\in C$,我们有$z_1\in C$。由(a)可得$[x,z_1\rangle\subset C^i$。从而$z\in C^i$。

2.17 例

下例表明无穷维凸集具备一个值得注意的结构。

(a) 假定V是具有可数基$B=\{x_1,x_2,\cdots\cdots\}$的线性空间,设$C=\mathrm{co}(\{0\}\cup B)$。我们有 $\mathrm{aff}\,(C)=V$,但$C^i=\phi$。事实上,设

$z=\sum_{i=1}^{k}\lambda_i x_i$ 是C的点，其中 $\lambda_i\geqslant 0$ $(1\leqslant i\leqslant k)$ 且$\sum_{i=1}^{k}\lambda_i=1$*（见定理2.3）。用$m$表示通过$z$和$x_N$的直线，其中$N>k$，则$m\cap C=[z,x_N]$，因此$z\in C^1$。

(b) 设V为例 (a) 中给出的空间，C是V中所有可以表示成$\sum_{i=1}^{k}\lambda_i x_i$的点之集合，其中$k\in N$，$\lambda_k>0$。$C$是$V$的正常凸子

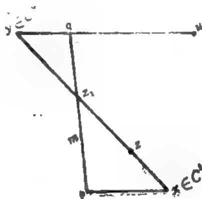


图8

集，但$C^0=V$。事实上，若$\sum_{i=1}^{m}\mu_i x_i\in V$，则对任意的$N>m$，$[x_N,z\rangle\subset C$。读者易证$C^1=\phi$及$C$相对于$V$的补集$D$亦为凸集且$D^0=V$，$D^1=\phi$。

---

* 译注：原文为$\sum_{i=1}^{k}\lambda_i\leqslant 1$。

凸代数体

2.18 定义

V 的代数内部非空的凸子集叫V 中的凸代数体。

2.19

设$C\subset V$是凸代数体。则对所有的 $a\in V$，集 $a+C$亦是凸代数体。若 $0\in C^i$，那么对任意的 $x\in V$，存在 $\delta>0$，使得 $[0,\delta x]\subset C$，因此$x\in(1/\delta)C$。我们给出下面定义。

定义. 设 C 为 V 中的凸代数体，$0\in C^i$. C 的度规（或 *Minkowski* 距离泛函）是用下式定义的函数p: $V\to\mathbf{R}$,

$$p(x)=\inf\{\lambda\geqslant 0\mid x\in\lambda C\}\quad(x\in V).$$

注.

若V是赋范线性空间，C 是V 中的单位球，则$p(x)$是 x 到原点 0 的距离.

2.20 定理

设 $C\subset V$ 为凸代数体，$0\in C^i$，C 的度规为p. 则对任意 x，$y\in V$，有

(a) $p(x)\geqslant 0$.

(b) $p(\lambda x)=\lambda p(x)\qquad\forall\lambda\geqslant 0$（$p$是正齐次的）.

(c) $p(x+y)\leqslant p(x)+p(y)$（$p$是次可加的）.

证明.

(a) 从p的定义立可得出.

(b) 是$\{\mu\geqslant 0\mid\lambda x\in\mu C\}=\lambda\{\mu\geqslant 0\mid x\in\mu C\}(\forall\lambda>0)$

的直接结果。

(c) 设 $p(x)=\alpha$，$p(y)=\beta$。对任意的 $\varepsilon>0$，$x\in(\alpha+\varepsilon)C$，$y\in(\beta+\varepsilon)C$，因此 $x+y\in(\alpha+\beta+2\varepsilon)C$（见§2.4，性质(b)）。由此得到

$(\forall\varepsilon>0)\quad p(x+y)\leqslant\alpha+\beta+2\varepsilon=p(x)+p(y)+2\varepsilon$，

于是 $p(x+y)\leqslant p(x)+p(y)$。

2.21

可用凸代数体的度规去描述它的代数内部和代数闭包的简单特征。

定理. 设 $C\subset V$ 为凸代数体，$0\in C^i$，p 为 C 的度规。则

$$C^i=(C^i)^i=(C^c)^i=\{x\in V\mid p(x)<1\},$$

$$C^c=(C^c)^c=(C^i)^c=\{x\in V\mid p(x)\leqslant 1\}.$$

证明. 首先注意到 $\forall x\in C$ 有 $p(x)\leqslant 1$，而 $x\in V$，$p(x)<1$ 则蕴涵 $x\in C$。

设 x，$y\in V$，$p(x)<1$。对充分小的 $\delta>0$，我们有 $p(x+\delta y)\leqslant p(x)+\delta p(y)<1$，* 因此 $[x,x+\delta y]\subset C$，即 $x\in C^i$。随后由 $[x,x+\delta y]\subset C^i$，得 $x\in(C^i)^i$。于是

$$p(x)<1\Longrightarrow x\in(C^i)^i \tag{19}$$

若 $p(x)=1$，对所有的 $z\in[0,x\rangle$，我们有 $p(z)<1$，因此 $[0,x\rangle\subset C$。从而 $x\in C^c$，乃至 $x\in(C^i)^c$。但 $x\notin C^i$，因为 $x\in C^i$ 蕴涵存在 $\lambda>1$，使 $\lambda x\in C$，由 $\lambda p(x)=p(\lambda x)\leqslant 1$，就有 $p(x)<1$。可见

$$p(x)=1\Rightarrow x\in(C^i)^c\setminus C^i \tag{20}$$

现在设 $x\in C^c$，$[z,x\rangle\subset C$。对 $\forall\lambda\in[0,1\rangle$，$\lambda x+(1-\lambda)z=z+\lambda(x-z)\in C$，因此 $p(z+\lambda(x-z))\leqslant 1$；在不等式 $p(x)\leqslant$

$p(z+\lambda(x-z))+(1-\lambda)p(x-z)\leqslant 1+(1-\lambda)p(x-z)$ 中，令$\lambda\uparrow 1$得到$p(x)\leqslant 1$。对集C^a用此结果，我们有

$$x\in(C^a)^a\Rightarrow p(x)\leqslant 1. \tag{21}$$

(19)式和(20)式蕴涵$C^i=(C^i)^i=\{x\in V|\ p(x)<1\}$，(19)、(20)、(21)式蕴涵$C^a=(C^a)^a=(C^i)^a=\{x\in V|\ p(x)\leqslant 1\}$。

最后，我们考虑集合$(C^a)^i$。如果$x\in(C^a)^i$，则存在$\lambda>1$，使$\lambda x\in C^a$，因此$\lambda p(x)=p(\lambda x)\leqslant 1$。从而$p(x)<1$。综上所述，我们得到$(C^a)^i=\{x\in V|\ p(x)<1\}$。

2.22 注

(a) 若$0\notin C^i$，但有凸集C的另外一点，譬于x_0，属于C^i。则平移C，使它包含0，利用$C-x_0$的度规p，并以$p(x-x_0)$代替$p(x)$，亦可得到类似上述的定理。

(b) 若$C\subset V$是凸的，$C^i=\phi$，则$(C^i)^i=\phi$。从而对任何凸集$C\subset V$有$(C^i)^i=C^i$成立（见§2.14，注(c)）。

(c) 凸代数体比代数内部为空集的凸集较少病态。例如，若$C^i\neq\phi$，则不可能有$C^a=V$而$C\neq V$（见§2.17）。事实上，$C^a=V$和$x_0\in C^i$蕴涵$V=\{x\in V|\ p(x-x_0)\leqslant 1\}$，其中$p$是$C-x_0$的度规。由此得知对$\forall x\in V$，$p(x-x_0)=0$，即$C=V$。

线性拓扑空间中的凸子集

下面，我们用E表示R上的线性拓扑空间，即E是具有拓扑的（在R上的）线性空间，使得加法（即$E\times E$到E中的映射：$(x,y)\to x+y$）和数乘（即$\mathrm{R}\times E$到E中的映射：$(\lambda,x)\to$

* 译注：原文为$\delta p(x)$。

λx）关于两个变量而言同时是连续的．这里$E\times E$和$\mathbf{R}\times E$上的拓扑是积拓扑．

为简单起见，我们假设E中不止含一个点，且E具有豪斯道夫（Hausdorff）拓扑

我们将用到下述简单性质．如果$U\subset E$是开的，则对每个非零$\lambda\in\mathbf{R}$、$a\in E$与$W\subset E$，λU，$a+U$和$U+W$亦是开集．

2.23 定理

设$C\subset E$是凸的．则

(a) $\text{int}(C)$和$\overline{C}$是凸的（见定理2.15）．

(b) 如果$x\in\text{int}(C)$，$y\in\overline{C}$，则$[x,y\rangle\subset\text{int}(C)$（见定理2.16）．

(c) $\text{int}(C)\subset C^i$。

(d) $C^i\subset\overline{C}$。

证明．(a) 设x_0，$y_0\in\overline{C}$，$\lambda\in\langle 0,1\rangle$，及$z_0=\lambda x_0+(1-\lambda)y_0$．假定$U$是$z_0$的一个邻域．由映射$(x,y)\to\lambda x+(1-\lambda)y$的连续性，存在$x_0$的一个邻域$U_1$和$y_0$的一个邻域$U_2$，使得$\lambda U_1+(1-\lambda)U_2\subset U$．因为$x_0$，$y_0\in\overline{C}$，存在$c_1\in U_1\cap C$及$c_2\in U_2\cap C$，则$\lambda c_1+(1-\lambda)c_2\in U\cap C$，因此$z_0\in\overline{C}$．$\overline{C}$是凸的．$\text{int}(C)$的凸性是(b)的一个结果．

(b) 首先注意，对$\forall a$，$b\in E$，$\forall\lambda\in\mathbf{R}$，$\lambda\neq 0$及$a$的任意邻域$U$，$b+\lambda(U-a)$是$b$的邻域．于是对$x_0\in\text{int}(C)$，$y_0\in\overline{C}$，$\lambda\in\langle 0,1\rangle$，及$z_0=\lambda x_0+(1-\lambda)y_0$．若$U$为$x_0$的邻域，且$U\subset C$，则

$$W=\frac{1}{1-\lambda}(z_0-\lambda U)$$

是 y_0 的一个邻域；由 $y_0 \in \overline{C}$，存在 $c \in W \cap C$。令

$$c = \frac{1}{1-\lambda}(z_0 - \lambda u),$$

于是，$z_0 = (1-\lambda)c + \lambda u$，其中 $u \in U$。则 $X = (1-\lambda)c + \lambda U$ 是 z_0 的一个邻域，根据 C 的凸性，我们有 $X \subset C$。

(c) 设 $x_0 \in \text{int}(C)$，$y_0 \in E$，U 是 x_0 的一个邻域，$U \subset C$。映射 $\lambda \to x_0 + \lambda y_0$ 在 0 的连续性蕴涵存在 $\delta > 0$，使对每个 $\lambda \in \mathbf{R}$，$|\lambda| < \delta$，有 $x_0 + \lambda y_0 \in U$，从而 $[x_0, x_0 + \delta y_0] \subset U \subset C$，所以 $x_0 \in C^i$。

(d) 我们有 $C \subset \overline{C}$。设 $y_0 \in C^a \setminus C$，$[x_0, y_0\rangle \subset C$ ，U 为 y_0 的邻域。由映射 $\lambda \to \lambda x_0 + (1-\lambda) y_0$ 在 0 的连续性，存在 $\delta > 0$，使得对 $\forall \lambda \in \mathbf{R}$，$|\lambda| \leqslant \delta$，有 $\lambda x_0 + (1-\lambda) y_0 \in U$。于是 $U \cup C \neq \phi$，因此 $y_0 \in \overline{C}$ 。

2.24 定理

若 $A \subset E$ 是开集，则 $\text{co}(A)$ 亦是开集。

证明. 我们有

$$A = \text{int}(A) \subset \text{int}(\text{co}(A)).$$

由定理2.23，$\text{int}(\text{co}(A))$ 是凸集，因此

$$\text{co}(A) \subset \text{int}(\text{co}(A)),$$

则 $\text{co}(A)$ 是开的。

注. 若 A 是闭集，一般说来 $\text{co}(A)$ 未必是闭，建议读者在 $\mathbf{R}^2$ 中构造一个反例。

2.25

设 $A \subset E$，E 中包含 A 的所有闭凸子集的交是 E 中包含 A 的

最小闭凸子集。

定义。设 $A\subset E$。A 的闭凸包 $\overline{\mathrm{co}}(A)$ 是 E 中包含 A 的最小闭凸子集。

由定理2.23(a)，我们有 $\overline{\mathrm{co}}\,(A)=\overline{\mathrm{co}(A)}$。

2.26 定义

E 中的凸体是 E 的一个凸子集 C 且 $\mathrm{int}(C)\neq\phi$。

由定理2.23(c)，凸体是凸代数体；其逆一般不真（见第三章习题 4）。

2.27 定理

设 $C\subset E$ 为凸体，则：

(a) $\mathrm{int}(C)=\mathrm{int}(\overline{C})=C^{i}$。

(b) $\overline{C}=\overline{\mathrm{int}(C)}=C^{a}$ 。

证明。(a) 设 $x\in\mathrm{int}(C)$，$y\in C^{i}$ 。存在点 z，使得 $y\in\langle z,x\rangle$ 且 $[z,y]\subset C$。由定理2.23(b)，有 $\langle z,x\rangle\subset\mathrm{int}(C)$，从而 $y\in\mathrm{int}\,(C)$。于是 $C^{i}\subset\mathrm{int}\,(C)$，再由定理 2.23 (c)，$C^{i}=\mathrm{int}(C)$。现设 $x\in\mathrm{int}\,(C)$，$y\in\mathrm{int}\,(\overline{C})$。存在点 z，使 $y\in\langle z,x\rangle$ 及 $[z,y]\subset\overline{C}$，于是 $z\in\overline{C}$。由定理2.23(b)，$y\in\mathrm{int}\,(C)$。故得 $\mathrm{int}(C)=\mathrm{int}(\overline{C})$。

(b) 设 $x\in\mathrm{int}\,(C)$，$y\in\overline{C}$，由定理 2.23 (b)，$[x,y\rangle\subset\mathrm{int}(C)$，因此 $y\in C^{a}$ 且 $y\in\overline{\mathrm{int}(C)}$，从而 $\overline{C}=\mathrm{int}\,(\overline{C})$，由定理2.23(d)，$\overline{C}=C^{a}$。

练　　习

V 和E 为本章中指定的空间。

1．设$A\subset V$是凸的，$x\in V\setminus A$．证明集

$$\bigcup_{a\in A}[x,a]$$

是凸的．

2．在 $\mathbf{R}^2$ 中，设$e_1=(1,0)$，$e_2=(0,1)$．
证明

$$\mathrm{co}(e_1,e_2,-e_1,-e_2)=\{(\xi_1,\xi_2)\mid|\xi_1|+|\xi_2|\leqslant1\}.$$

3．设d 为 $\mathbf{R}^n$ 上的欧几里得度量，若 $C\subset\mathbf{R}^n$ 是凸的，证明：对$\forall\varepsilon>0$，

$$\{x\in\mathbf{R}^n\mid d(x,C)\leqslant\varepsilon\}$$

是凸的．

4．见§2.9.证明V中凸多胞形的顶点集是唯一确定的．

5．证明

$$C=\{(\xi_1,\xi_2,\cdots,\xi_n)\mid\xi_i\geqslant0\ (1\leqslant i\leqslant n)，且\sum_{i=1}^{n}\xi_i=1\}$$

是 $\mathbf{R}^n$ 中的凸多胞形．求出 C 的顶点．

6．证明凸多胞形的极端点是它的顶点（见§2.10）．

7．见§2.17例（b)．证明 D 是凸的，且 $C^i=D^i=\phi$，$D^a=V$．

8．设p：$V\to\mathbf{R}$ 具有定理 2.20 提及的性质(a)，(b)和(c)．设 $A=\{x\in V\mid p(x)<1\}$，$B=\{x\in V\mid p(x)\leqslant1\}$．证明：

(a) A和B是凸集，且$0\in A^i\cap B^i$．

(b) $A^i = A$, $B^a = B$.

(c) p 是A和B的度规。

9. 设$C \subset V$ 为凸代数体，$0 \in C^i$，p 为C的度规。证明：对所有 $x \in V$

$$p(x) = 0 \Longleftrightarrow x = 0 \text{ 或者始点在0 且通过 } x \text{ 的射线含于 } C \text{ 中。}$$

10. 设$C \subset E$是凸的且$0 \in C^i$，p为C的度规，证明C是凸体当且仅当p是连续函数。

利用这个结果，再次证明（见定理2.27）若$C \subset E$是凸体，则$\operatorname{int}(C) = C^i$ 且$\overline{C} = C^a$。

11. 设 C_1，C_2 为E中的凸体。证明 $\overline{C}_1 = \overline{C}_2$，当且仅当 $\operatorname{int}(C_1) = \operatorname{int}(C_2)$。

12. 设 $A \subset E$ 是闭的且非凸。证明存在 x，$y \in A$，使得 $A \cap \langle x, y \rangle = \phi$。

13. 设$C_1, C_2, \cdots, C_n \subset E$是凸的，使

$$\bigcap_{i=1}^{n} \operatorname{int}(C_i) \neq \phi,$$

证明

$$\overline{\bigcap \operatorname{int}(C_i)} = \bigcap \overline{C}_i.$$

注　　释

关于代数内部与代数闭包更详细的论述可见下述文献

J.Blair and R.Fourneau, *Étude géométrique des espaces vectoriels*, Berlin, Springer. (part 1, *Lecture Notes in Mathematics* no. 489, 1975 and part Ⅱ, *Lecture Notes in Mathematics* no. 802 1980).

第三章　分离定理

线性空间中的分离

我们用V表示至少含两点的$\mathbf{R}$上的线性空间。

3.1　定理

设A，$B\subset V$是凸的且$A\cap B=\phi$。则存在V的凸子集C,D，满足

$$C\cup D=V,\ C\cap D=\phi,\ C\supset A,\ D\supset B.$$

证明. 设M是所有V的凸子集序对(P,Q)构成的集合，此处$P\cap Q=\phi$，$P\supset A$，$Q\supset B$。定义M的一个偏序'$\leqslant$'如下：如果$P_1\subset P_2$，$Q_1\subset Q_2$。就说$(P_1,Q_1)\leqslant(P_2,Q_2)$。$M$的每一线性有序子集$\{(P_\alpha,Q_\alpha)\}$在$M$中有一个上界，即$(\bigcup_\alpha P_\alpha,\bigcup_\alpha Q_\alpha)$。由曹恩($Zorn$)引理，$M$有极大元$(C,D)$，且$C\supset A,D\supset B$，$C\cap D=\phi$。读者易证$C\cup D=V$。(假若存在$x\in V,x\notin C\cup D$，那么就有$\mathrm{co}\,(\{x\}\cup C)\cap D=\phi$或$\mathrm{co}\,(\{x\}\cup D)\cap C=\phi$，即序对$(C,D)$不是极大元)。

3.2

我们说V的线性子空间L有余维数1，如果存在$p\in V$，

$p \notin L$，使

$$V = L + \mathrm{R}p \quad (= \{m + \lambda p \mid m \in L, \lambda \in \mathrm{R}\})。$$

V中的超平面是指V中形如

$$H = a + L$$

的子集H，其中$a \in V$，L是V的余维数为1的线性子空间。

设$H = a + L$是V中的一个超平面。我们定义f：$V \to \mathrm{R}$如下：若$x \in V$，$x = m + \lambda p$，则$f(x) = \lambda$。此f必是线性函数，且$f(p) = 1$，$H = f^{-1}(f(a))$ $(= \{x \in V \mid f(x) = f(a)\})$。

反之，若f：$V \to \mathrm{R}$是一个不恒为0的线性函数，$a \in \mathrm{R}$，那么$f^{-1}(a)$是一个超平面。事实上，存在$p \in V$，满足$f(p) = 1$。如果$x \in V$，则$f(x - f(x)p) = 0$，因此，$x - f(x)p \in f^{-1}(0)$，于是$x \in f^{-1}(0) + f(x)p$*。由此得出$V = f^{-1}(0) + \mathrm{R}p$，从而$f^{-1}(0)$（这是$V$的线性子空间）具有余维数1。最后，$f^{-1}(a) = ap + f^{-1}(0)$，即$f^{-1}(a)$是一个超平面。

这样，我们已证明了下面定理：

H是V中的超平面，当且仅当存在线性函数f：$V \to \mathrm{R}$，$f \neq 0$和$a \in \mathrm{R}$，使得

$$H = f^{-1}(a)。$$

.3 定理

设C, D为V中非空凸子集，$C \cap D = \phi$，$C \cup D = V$。令$H = C^c \cap D^c$，则或者H是超平面；或者$H = V$。

证明. 首先我们注意，如果$x \in V$，$x \notin C^c$，则$x \in D^c$（这个简单的证明留给读者）。从而$V = H \cup C^c \cup D^c$。

---

* 译注：原文 $x = f^{-1}(0) + f(x)p$，有错。

设 $x,y\in H$。由定理2.15，H是凸的，因此 $[x,y]\subset H$。我们将证明通过 x 和 y 的直线 m 含于 H。假设不然，则 m 上有点 $z\notin[x,y]$且 $z\notin H$。这蕴涵 $y\in\langle x,z\rangle$ 或 $x\in\langle y,z\rangle$，譬如 $y\in\langle x,z\rangle$。就有 $z\in C^i\cup D^i$，不妨设 $z\in C^i$。由定理2.16，这将推出 $y\in C^i$，然而 $y\in D^i$，此与假定矛盾。

这就得出H是一个仿射集。

在继续证明前，我们注意，若 $x\in C$，$y\in D$，则 $[x,y]\cap H\neq\phi$（这个简单的证明留给读者）。

现在，假定$H\neq V$，$a\in V\setminus H$。设 $h\in H$；我们将证明V的线性子空间$H\text{-}h$ 有余维数 1（这蕴涵H是一个超平面）。我们有 $a\in C^i\cap D^i$，设 $a\in C^i$。我们也有$2h-a\notin H$，因此 $2h-a\in C^i\cup D^i$。若$2h-a\in C^i$，由 $a\in C^i$ 将导致 $h\in C^i$（根据C^i 的凸性），出现矛盾。由此 $2h-a\in D^i$（见图 9）。

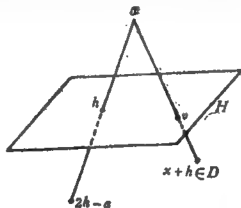


图 9

设 $x\in V$，则 $x+h\in C\cup D$。若 $x+h\in C$，则$[x+h,2h-a]\cap H\neq\phi$，于是，存在 $\lambda\in(0,1]$ 及 $v\in H$，使 $\lambda(x+h)+$

$(1-\lambda)(2h-a)=v$，因此

$$x=\left(\frac{1}{\lambda}-1\right)(a-h)+\frac{1}{\lambda}(v-h).$$

若 $x+h\in D$，则 $[x+h,a]\cap H\neq\phi$，于是存在 $\lambda\in\langle 0,1]$ 及 $v\in H$，使 $\lambda(x+h)+(1-\lambda)a=v$，因此

$$x=\left(\frac{1}{\lambda}-1\right)(h-a)+\frac{1}{\lambda}(v-h).$$

在两种情况下，都存在 $\sigma\in\mathbf{R}$ 和 $y\in H-h$ 使 $x=\sigma(a-h)+y$。由此得出 $V=H-h+\mathbf{R}(a-h)$，因此 $H-h$ 有余维数 1。

3.4

为简便起见，我们常写 $f(A)<\alpha$ 代替 $(\forall x\in A)f(x)<\alpha$。

设 A，$B\subset V$，$H=f^{-1}(\alpha)$ 是 V 中一个超平面。

(a) 我们说 H 分离 A 和 B，如果 $f(A)\leqslant\alpha$，$f(B)\geqslant\alpha$，或 $f(A)\geqslant\alpha$ $f(B)\leqslant\alpha$。

(b) 我们说 H 真分离 A 和 B，如果 H 分离 A 和 B 且 A，B 两者都无点被含于 H（即存在 $x\in A\cup B$，使 $f(x)\neq\alpha$）。

(c) 我们说 H 严格分离 A 和 B，如果 $f(A)<\alpha$，$f(B)>\alpha$ 或 $f(A)>\alpha$，$f(B)<\alpha$。

3.5 定理（分离定理）

设 A，B 为 V 中凸集，$A\neq\phi$，$B^i\neq\phi$，$A\cap B^i=\phi$。则在 V 中存在一个超平面真分离 A 和 B。

证明. 由定理3.1，存在 V 的凸子集 C，D，使 $C\cup D=V$，$C\cap D=\phi$，$C\supset A$，$D\supset B^i$。我们把 $B^i\cap C^a=\phi$ 留给读者证明。由 $B^i\neq\phi$，得出 $C^a\cap D^a\neq V$。由定理3.3，$H=C^a\cap D^a$ 是一个超

平面。设$H=f^{-1}(\alpha)$。由$B^i\cap H=\phi$，有$b\in B^i$，使$f(b)\neq\alpha$，不妨设$f(b)>\alpha$。假设存在$x\in B$，使$f(x)<\alpha$。因为f是线性的，这蕴涵着存在$z\in\langle x,b\rangle$，使$f(z)=\alpha$，因此$z\in H$。但由定理2.16，$\langle x,b]\subset B^i$，所以$z\in B^i$，这与$H\cap B^i=\phi$矛盾。从而得$f(B)\geqslant\alpha$。

现在假设存在$x\in A$，使$f(x)>\alpha$。因为$[x,b]\cap H\neq\phi$，这蕴涵着存在$z\in\langle x,b\rangle$，使$f(z)=\alpha$。但由于$f(x)>\alpha, f(b)>\alpha$，则对任意的$y\in\langle x,b\rangle$，恒有$f(y)>\alpha$，矛盾。从而导出$f(A)\leqslant\alpha$。

由此得到H分离A和B。因为存在$b\in B$，使$f(b)\neq\alpha$，故为真分离。

注。 条件$B^i\neq\phi$是很重要的。例如，再次考虑§2.17例（b）中的集合C，$C^i=\phi$，$C^a=V$。若存在超平面$H=f^{-1}(\alpha)$，使$f(C)\leqslant\alpha$，因为$C^a=V$，我们恒有$f(V)\leqslant\alpha$，这不可能，由此得出集C与点$x\notin C$不能被超平面分离。

线性拓扑空间中的分离

我们用E表示R上的线性拓扑空间，E不只含一个点且具有Hausdorff 拓扑。

3.6 定义

设$A\subset E$，$a\in A$。A被称为相对于a是星形的，当对任意的$x\in A$，$[a,x]\subset A$。

定理. 设开集$U\subset E$，$a\in U$。则存在E的开子集S，使得$S\subset U$，且S相对于a是星形的（换言之，E中每点的邻域系存在由星形集组成的基）。

证明。因为$W=U-a$是点0的邻域，映射$(\lambda,x)\to\lambda x$在$(0,0)$的连续性推出存在$\varepsilon>0$，及$0\in E$的邻域U_0，使得对所有的$\lambda\in\mathbf{R}$，$|\lambda|<\varepsilon$，有$\lambda U_0\subset W$。令

$$X=\bigcup_{|\lambda|<\varepsilon}\lambda U_0$$

则X是0的邻域，相对于0是星形的，且$X\subset W$。从而$X+a$是a的邻域，相对于a星形的，且$X+a\subset U$。

3.7 定理

设$H=f^{-1}(\alpha)$为E中的超平面，令$A=\{x\in E\mid f(x)>\alpha\}$，则下列条件等价：

(a) H是闭的。

(b) $\mathrm{int}(A)\neq\phi$。

(c) f是连续的。

证明。(a)$\Rightarrow$(b)：设$x_0\in A$。因为H是闭的，存在x_0的邻域U，使$U\cap H=\phi$。由§3.6，存在x_0的邻域S，相对于x_0是星形的，且$S\subset U$。由是$f(S)>\alpha$，故$S\subset A$。A是开集。

(b)$\Rightarrow$(c)：设$x_0\in\mathrm{int}(A)$，$x\in A$且$x\neq x_0$。因为$f(x_0)>\alpha$，$f(x)>\alpha$，于是存在$z\in A$，使$x\in\langle z,x_0\rangle$。由定理2.23 (b)，利用A的凸性，即得$x\in\mathrm{int}(A)$。A是开的。设$(\sigma,\tau)\subset\mathbf{R}$，我们证明$f^{-1}(\sigma,\tau)$是开的（这推出$f$的连续性）。设$b\in E$使$f(b)=1$。我们有

$$\begin{aligned}f(x)\in\langle\sigma,\tau\rangle&\Longleftrightarrow\sigma<f(x)<\tau\\&\Longleftrightarrow f(x+(\alpha-\sigma)b)=f(x)+\alpha-\sigma>\alpha\text{ 且 }\\&\qquad f((\alpha+\tau)b-x)=\alpha+\tau-f(x)>\alpha\\&\Longleftrightarrow x+(\alpha-\sigma)b\in A\text{ 且 }(\alpha+\tau)b-x\in A\\&\Longleftrightarrow x\in A-(\alpha-\sigma)b\text{ 且 }x\in-A+(\alpha+\tau)b\end{aligned}$$

从而得$f^{-1}(\sigma,\tau)=[A-(\alpha-\sigma)b]\cap[-A+(\alpha+\tau)b]$。因$A$是开的，所以$-A$亦是开集。故$f^{-1}(\sigma,\tau)$是开的。

(c)$\Rightarrow$(a)：因为$\{\alpha\}$是闭的（在R中），所以$f^{-1}(\alpha)$亦闭。

注。

(a) 若$x_0\in H$，则$\{x\in E|f(x)<\alpha\}=2x_0-A$。由此推出 $\text{int}(A)\neq\phi$ 当且仅当 $\text{int}\{x\in E|f(x)<\alpha\}\neq\phi$。

(b) 设$H=f^{-1}(\alpha)$为E中超平面，则$\overline{H}$为仿射集（这留给读者证），有

$$\text{co dim}\ (\overline{H})\leqslant\text{co dim}\ (H)=1$$

其中 $\text{co dim}\ (H)$ 为H的余维数。若H非闭，则$\overline{H}\neq H$，因此 $\text{co dim}\ (\overline{H})=0$，故$\overline{H}=E$。换言之，若$H$非闭，$H$在$E$中必稠密。

(c) 每一个线性函数 f：$R^n\to R$是连续的（因此R^n中每个超平面是闭的）。我们给出一个非连续的线性函数 f：$E\to R$的例子。设$[a,b]\subset R$，E是所有可微函数 x：$[a,b]\to R$ 组成的线性赋范空间，其范数为

$$\|x\|=\sup_{a\leqslant t\leqslant b}|x(t)|\qquad(x\in E)$$

读者可证用 $f(x)=x'(a)$ 定义的函数 f：$E\to R$是线性的，但不连续。

3.8 定理（分离定理）

设$A,B\subset E$是凸的，$A\neq\phi$，$\text{int}(B)\neq\phi$，$A\cap\text{int}(B)=\phi$。则存在E中闭超平面真分离A和B。

证明。由定理2.27(a)，$\text{int}(B)=B^i$。按照定理3.5，存在E中超平面$H=f^{-1}(\alpha)$真分离A和B。不失一般性，假定$f(B)\geqslant\alpha$，从而$f(\text{int}(B))>\alpha$（证明留给读者）。根据定理3.7，H是闭的。

Hahn—Banach 定 理

3.9

定理3.8是所调 Hahn—Banach 定理的一种形式。它最为

人熟知的表述是（用于泛函分析中）涉及到定义在E的线性空间上的（连续）线性泛函的延拓，其另一几何形式如下

定理

设$C\subset E$为凸体，M为E中仿射集，$M\neq\phi$，且$M\cap\operatorname{int}(C)=\phi$。则在$E$中存在一个闭超平面包含$M$，且真分离$C$和$M$。

证明．由定理3.8，有个E中闭超平面$H=f^{-1}(\alpha)$真分离C和M。设$m\in M$，$f(m)=\beta$。读者易证闭超平面$f^{-1}(\beta)$包含M，且真分离C和M。

3.10 应用：支撑超平面

设$A\subset E$，$x\in\operatorname{bd}(A)$。E中超平面$H=f^{-1}(\alpha)$称为A在x的支撑超平面，如果

(i) $x\in H$,

(ii) $f(A)\geqslant\alpha$或$f(A)\leqslant\alpha$。

（A被认为位于由H确定的两个半空间之一中），A的支撑超平面H称为是非平凡的，若A不含于H。

定理．

设$C\subset E$为凸体，$x\in\operatorname{bd}(C)$，则C在X存在非平凡的闭支撑超平面。

证明．这是§3.9的直接结果．可直接证明如下：

由定理3.8，存在E中闭超平面$H=f^{-1}(\alpha)$真分离C和$\{x\}$。假定$f(C)\leqslant\alpha$，$f(x)\geqslant\alpha$。由$x\in\operatorname{bd}(C)$，又有$f(x)\leqslant\alpha$，于是$f(x)=\alpha$，即$x\in H$。又因是真分离，有$C\cup\{x\}\not\subset H$，从而$C\not\subset H$。所以H是C在x非平凡的闭支撑超平面。

3.11

在凸分析中，Hahn—Banach 定理的大多数应用涉及到线性拓扑空间的两种特殊情形，即

(a) 局部凸空间. 这些线性拓扑空间中每点的邻域系有一个由凸集组成的基. 鉴于读者也许不熟悉这种空间的理论（在应用中，这种空间不太常见），我们将限于线性赋范空间. 在这种空间中，每点 a 的邻域系有一个由球$\{x \mid \|x-a\|<\varepsilon\}$组成的基，其中$\varepsilon>0$. 如果 A，B 是这种空间中的两个子集，我们定义：

$$d(A,B):=\inf\{\|a-b\| \mid a\in A,\ b\in B\}$$

(b) 有限维空间. 这种空间与赋有某种范拓扑的R^n（对R^n而言所有范数是等价的）是拓扑同构的（线性同胚）R^n 的凸性将在下一章研究.

线性赋范空间中的定理

3.12 定理

设E是线性赋范空间，C为E 中非空闭凸集. $a\notin C$，则存在E中闭超平面严格分离C和a.

证明. 因为C是闭的，我们有

$$\sigma:=d(a,C)=\inf_{x\in C}\|x-a\|>0$$

令 $C_{\frac{1}{2}\sigma}=\{x\in E \mid d(x,C)<\frac{1}{2}\sigma\}$，$B=\{x\in E \mid \|x-a\|<$

$\frac{1}{2}\sigma$)。$C_{\frac{1}{2}\sigma}$与B是凸集，且有$C_{\frac{1}{2}\sigma}\cap B=\phi$。由定理3.8，存在$E$中闭超平面分离$C_{\frac{1}{2}\sigma}$和$B$。易证它严格分离$C$和$a$。

3.13 应用：凸锥

在凸分析应用中，例如对优化条件的开发中，凸锥的概念及性质是非常有用的。

定义. 设V为R上线性空间。$K\subset V$称为锥，如果对任何$x\in K$，$\lambda>0$，$\lambda x\in K$。

所谓K的顶点0，可以属于K，也可以不属于K。许多作者通过在引入锥的定义时，令$\lambda\geqslant 0$，以包含顶点0。

读者易证$K\subset V$为凸锥当且仅当对所有的$x,y\in K,\lambda>0$，有

$$x+y\in K,\ \lambda x\in K.$$

定义. 设E是线性赋范空间，E'为对偶，即E上的所有线性连续泛函$E\to$R组成的线性空间。若$x\in E$，$u\in E'$，我们写$\langle x|u\rangle$代替$u(x)$。设$K\subset E$是一个锥。

(a) K的极用K^0表示，

$$K^0:=\{u\in E'|\ \ \langle K|u\rangle\leqslant 0\}$$

其中我们写$\langle K|u\rangle\leqslant 0$代替$(\forall x\in K)\langle x|u\rangle\leqslant 0$。

(b) K的二次极，用K^{00}表示。

$$K^{00}:=\{x\in E|\ \langle x|K^0\rangle\leqslant 0\}$$

读者易证K^0和K^{00}是包含点0的凸锥。我们有$K\subset K^{00}$。下面证明K^{00}是闭的。事实上，设(x_n)为K^{00}中序列，且$x_n\to x$。若$u\in K^0$，则对所有的$n\in$ N，$\langle x_n|u\rangle\leqslant 0$，因此，$\langle x|u\rangle\leqslant 0$

（由于 u 的连续性），于是 $x \in K^{00}$。得出$\overline{K} \subset K^{00}$。

定理. 若$K \subset E$为非空凸锥，则$K^{00} = \overline{K}$

证明：已证$\overline{K} \subset K^{00}$。读者易证$\overline{K}$是包含点 0 的凸锥。设 $a \bar{\in} \overline{K}$*，由定理3.12，存在 $u_0 \in E'$及$\alpha \in \mathbf{R}$，使

$$\langle \overline{K} \mid u_0 \rangle < \alpha \tag{22}$$

和

$$\langle a \mid u_0 \rangle > \alpha.$$

我们有 $0 \in \overline{K}$，因此 $\alpha > 0$。所以

$$\langle a \mid u_0 \rangle > 0 \tag{23}$$

因为对所有$\lambda > 0$，$\lambda \overline{K} = \overline{K}$，（22）式蕴涵$\langle \overline{K} \mid u_0 \rangle \leqslant 0$，即$u_0 \in K^0$。所以，从（23）式得出 $a \bar{\in} K^{00}$。

我们得$K^{00} \subset \overline{K}$，从而证明了$K^{00} = \overline{K}$。

练　　习

1　设E为线性拓扑空间。A，$B \subset E$是非空的开凸集，使$A \cap B = \phi$。证明存在E中闭超平面严格分离A和B。

2　设E为线性赋范空间。$A \subset E$ 是非空的闭凸集，$B \subset E$是非空的紧凸集，使$A \cap B = \phi$。证明存在E 中闭超平面严格分离A和B（见练习 1 和定理3.12）。

3　设E为线性拓扑空间。$A \subset E$是凸的且$\text{int}(A) \neq \phi$。证明 $x \in \text{int}(A)$当且仅当对每一个包含 x 的超平面H，A 中至少存在两个点能被H严格分离。

4　给出是凸代数体而不是凸体的例子（见 § 2 26）。

---

* 译注：原文误为$a \bar{\in} K$。

（提示：研究第三章习题 8 和习题10）。

5　设K是线性赋范空间中的开凸锥。证明$K+\overline{K}=K$。

6　设K_1，K_2为线性赋范空间中非空的锥．证明

(a)　$(K_1+K_2)^0=K_1^0\cap K_2^0=(K_1\cup K_2)^0$

(b)　$(K_1\cap K_2)^0\supset K_1^0+K_2^0=\mathrm{co}(K_1^0\cup K_2^0)$（见第四章练习13）。

第四章　$\mathbf{R}^n$中的凸子集

4.1

本章我们研究$\mathbf{R}^n$中的凸子集（$n\geqslant 1$），或者说有限维线性空间的凸子集，就本质而言，两者是相同的。

某些经典定理

4.2　定理（Caratheodory定理）

设$A\subset\mathbf{R}^n$。

(a) 对于每点$x\in\mathrm{co}(A)$，存在A的$n+1$个点，使得x是这些点的凸组合。

(b) 若$\dim(A)=k$，则$\mathrm{co}(A)$是所有顶点属于A的k－单纯形的并。

证明。设$x\in\mathrm{co}(A)$。由定理2.3，存在x_1，x_2，…，$x_p\in A$和λ_1，λ_2，…，$\lambda_p\in\mathbf{R}$，$\lambda_i\geqslant 0$（$1\leqslant i\leqslant p$）使得

$$\sum_{i=1}^{p}\lambda_i=1 \text{ 且 } x=\sum_{i=1}^{p}\lambda_i x_i .$$

若x_1，x_2，…，x_p是仿射相关的，则存在不全为0的μ_1，μ_2，…，$\mu_p\in\mathbf{R}$，使

$$\mu_1x_1+\mu_2x_2+\cdots+\mu_px_p=0$$

且

$$\mu_1+\mu_2+\cdots+\mu_p=0$$

（见 § 2.11）．对所有的$\rho\in\mathbf{R}$，我们有

$$x=(\lambda_1-\rho\mu_1)x_1+\cdots+(\lambda_p-\rho\mu_p)x_p. \qquad (24)$$

读者易证集$\{\sigma\in\mathbf{R}|\ \sigma\mu_i\leqslant\lambda_i(1\leqslant i\leqslant p)\}$是一个闭区间$[\alpha,\ \beta]$（注意至少存在一个$i$，使$\mu_i\neq 0$），在(24)式中，令$\rho=\alpha$．得到

$$x=(\lambda_1-\alpha\mu_1)x_1+\cdots+(\lambda_p-\alpha\mu_p)x_p$$

其中

$$\sum_{i=1}^{p}(\lambda_i-\alpha\mu_i)=1,\ \lambda_i-\alpha\mu_i\geqslant 0\ (1\leqslant i\leqslant p).$$

且至少存在一个i，使

$$\lambda_i-\alpha\mu_i=0.$$

这就推知x是A中$p-1$个点的凸组合，不妨假定为x_2，x_3，$\cdots$，x_p．如x_2，$x_3\cdots$，x_p还仿射相关，可重复上述过程．若$p>n+2$，这是一定可能的，经有限步后，可得x为A的q个仿射独立点y_1，y_2，$\cdots$，y_q的凸组合，其中$q\leqslant n+1$．

(a) 若$q=n+1$，立得所述结果．若$q<n+1$，则从A中任选其他点y_{q+1}，$\cdots$，y_{n+1}，将项$0\cdot y_{q+1}+0\cdot y_{q+2}+\cdots+0\cdot y_{n+1}$加到由$y_1$，$y_2$，$\cdots$，$y_q$给出的式中，即得欲求的$x$的表达式．

(b) 若$\dim(A)=k$，则$q\leqslant k+1$．若$q=k+1$，则x含于顶点属于A的k-单纯形$cc(y_1,\ y_2,\ \cdots,\ y_{k+1})$．若$q<k+1$，可从$A$中选择其他点$y_{q+1}$，$y_{q+2}$，$\cdots$，$y_{k+1}$，使$y_1$，$y_2$，$\cdots$，$y_{k+1}$为仿射独立．利用(a)的论证，即得所述结果．

注．

(a) 在R^n中，Caratheodory定理是定理2.3的一种改进。

(b) 若$A\subset R^n$至多有n个分支，且$\dim(A)=n$，可以证明co(A)是顶点属于A的所有$(n-1)$-单纯形之并。

读者易证（用一个合适的例子）在这个命题中，数$n-1$不能改进.

4.3 应用

由定理2.24，若A为开集，则co(A)亦开，但A为闭集，则co(A)未必是闭。然而，在R^n中，只要A是紧集，则co(A)也是紧集。

定理。若$A\subset R^n$是紧集，则co(A)亦是紧集。

证明。由Caratheodory定理，co(A)中的点可被表为A中$n+1$个点的凸组合。用

$$f(x_1,\ \cdots,\ x_{n+1},\ \lambda_1,\ \cdots\lambda_{n+1})=\sum_{i=1}^{n+1}\lambda_i x_i$$

定义的映射f：$R^{n(n+1)}\times R^{n+1}\longrightarrow R^n$（$x_i\in R^n$，$\lambda_i\in R$（$1\leqslant i\leqslant n+1$））是连续的。设

$$L=\{(\lambda_1,\ \cdots,\ \lambda_{n+1})\in R^{n+1}\mid \lambda_i\geqslant 0\ (1\leqslant i\leqslant n+1),$$

$\sum_{i=1}^{n+1}\lambda_i=1\}$则co($A$)$=f(A^{n+1}\times L)$。读者易证$L$是紧的。从而作为紧集的笛卡尔乘积$A^{n+1}\times L$是紧的。因此它在$f$下的象亦是紧集。

4.4 定理（Helly定理）

对于R^n中至少$n+1$个凸集组成的有限集族，若每个由$n+1$个集合组成的子族有非空的交，则整个集族亦有非空的交。

证明. 设集族由N个集组成。对N采用归纳法证明。当$N = n+1$时，定理显然是真。设$N = k$时（这里$k \geqslant n+1$）所述结果成立，考虑$N = k+1$的情形，假定集族是$\{C_1, C_2, \cdots, C_{k+1}\}$。由归纳假设，对每个$i$，$1 \leqslant i \leqslant k+1$，存在$x_i \in R^n$，使得

$$x_i \in \bigcap_{\substack{j=1 \\ j \neq i}}^{k+1} C_j$$

因为$k > n$，点$x_1, x_2, \cdots, x_{k+1}$是仿射相关的。故存在不全为0的$\mu_1, \mu_2, \cdots, \mu_{k+1} \in R$，满足

$$\sum_{i=1}^{k+1} \mu_i x_i = 0, \quad \sum_{i=1}^{k+1} \mu_i = 0.$$

不失一般性，我们假定$\mu_1 \geqslant 0, \cdots, \mu_s \geqslant 0, \mu_{s+1} < 0, \cdots, \mu_{k+1} < 0$。置

$$y = \sum_{i=1}^{s} \mu_i x_i \Big/ \sum_{i=1}^{s} \mu_i$$

那么，我们也有

$$y = \sum_{i=s+1}^{k+1} (-\mu_i) x_i \Big/ \sum_{i=s+1}^{k+1} (-\mu_i)$$

从而，y既属于$\bigcap_{i=s+1}^{k+1} C_i$也属于$\bigcap_{i=1}^{s} C_i$，即$y$是所有集合$C_i (1 \leqslant i \leqslant k+1)$的公共点。

注. 对由R^n中的紧凸集组成的任意集族(不必是有限的)有类似定理.

4.5

$\mathbf{R}^n$中的超平面是$\mathbf{R}^n$的形如$\{x\in\mathbf{R}^n|(\alpha|x)=\beta\}$的子集，这里$\alpha\in\mathbf{R}^n\backslash\{0\}$，$\beta\in\mathbf{R}$，$(\alpha|x)=\alpha_1\xi_1+\alpha_2\xi_2+\cdots+\alpha_n\xi_n$是向量$\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_n)$和$x=(\xi_1,\xi_2,\cdots,\xi_n)$的内积（见§3.2）。

下述定理是Helly定理的一个简单应用。

定理（kirchberger定理）。设A，B是$\mathbf{R}^n$中的有限集，假定对每个$A\cup B$的子集C，它至多由$n+2$个点组成，且存在严格分离$C\cap A$和$C\cap B$的超平面，则存在超平面严格分离A和B。

证明。假设$A\cup B$不止含$n+2$个点。对每个$x\in\mathbf{R}^n$，我们定义

$$A(x)=\{(y,\beta)\in\mathbf{R}^n\times\mathbf{R}|(x|y)>\beta\}$$
$$B(x)=\{(y,\beta)\in\mathbf{R}^n\times\mathbf{R}|(x|y)<\beta\}$$

$A(x)$和$B(x)$是$\mathbf{R}^n\times\mathbf{R}=\mathbf{R}^{n+1}$的凸子集。考察由所有的$A(x)$，$x\in A$和所有的$B(x)$，$x\in B$组成的有限集族$P$。由题设可知，对$P$的每$n+2$个元素$A(x_1)$，$A(x_2)$，$\cdots$，$A(x_p)$，$B(x_{p+1})$，$\cdots$，$B(x_{n+2})$存在$(y,\beta)\in\mathbf{R}^n\times\mathbf{R}$，使得

$$\begin{cases}(y|x_i)>\beta,\ \text{因此}(y,\beta)\in A(x_i)\ (1\leqslant i\leqslant p)\\(y|x_i)<\beta,\ \text{因此}(y,\beta)\in B(x_i)\ (p+1\leqslant i\leqslant n+2)\end{cases}$$

故(y,β)是这$n+2$个元素之交。由Helly定理，P之所有元素必有公共点，定理得证。

注 读者易证在这个定理中，数$n+2$不能改进。

4.6

令Ω_{cc}是$\mathbf{R}^n$的所有非空紧凸子集组成的集族。设A，$B\in$

Ω_{CC}。A在B上的超出量定义为

$$e(A,\ B):=\sup\{d(x,\ B)|\ x\in A\}$$

这里d是R^n上的欧几里得度量。A和B的Hausdorff距离定义为

$$h(A,\ B):=\max\{e(A,\ B),\ e(B,\ A)\}$$

具有Hausdorff距离的Ω_{CC}是一个度量空间，证明留给读者。

设$R>0$，令$\Omega_{CC}(R)$是球$\{x\in R^n|\ d(x,\ 0)\leqslant R\}$的所有非空紧凸子集全体。

定理（Blaschke 收敛定理）．具有 Hausdorff 距离的$\Omega_{CC}(R)$是紧度量空间。

证明．只要证明$\Omega_{CC}(R)$是序列紧就够了．设P为n维立方体

$$\{(\xi_1,\ \xi_2,\ \cdots,\ \xi_n)\in R^n|\ 0\leqslant|\xi_i|\leqslant R(1\leqslant i\leqslant n)\}$$

通过把P的每边等分成2^k份，我们得到P的一个划分，将P分成一些边长为$R/2^{k-1}$的全等闭立方体。用P^k表示这些闭立方体的集族。任一$A\in\Omega_{CC}(R)$均含于P，P^k中所有与A相交的立方体之并称为A的一个k-复盖。若A和B有相同的k-复盖，则有$(A,\ B)<R\sqrt{n}/2^{k-1}$。

设(C_i)为$\Omega_{CC}(R)$中一个序列。因为不同的1－复盖的个数是有限的，则存在(C_i)的子序列$(C_j^{(1)})$，其中元素具有相同的1－复盖。同理，存在$(C_j^{(1)})$的子序列$(C_j^{(2)})$，其元素有相同的2－复盖。继续使用这种方法，对每一个$k\in N$，我们得到一列有下述性质的$(C_j^{(k)})$：

(a) $(C_j^{(k)})$中所有元素有相同的k-复盖，因此对所有的j，$m\in N$，我们有$h(C_j^{(k)},\ C_m^{(k)})<R\sqrt{n}/2^{k-1}$

(b) $(C_j^{(k)})$是$(C_j^{(k-1)})$的子序列，因此对所有的j，$m\in N$及所有的p，$q\in N$，且$p>q$，我们有$h(C_j^{(p)},\ C_m^{(q)})$

$< R\sqrt{n}/2^{q-1}$。

由此得知，对角线序列 (D_j) 是一个柯西序列，这里 $D_j = C^{(j)}_j$。令

$$A_m = \overline{\text{co}}\ (\bigcup_{j \geq m} D_j)$$

(A_m) 是非增的非空紧凸集列。而

$$A^{*} = \bigcap_{m \geq 1} A_m$$

为非空紧凸集，所以 $A \in \Omega_{CC}(\mathbf{R})$。我们指出有 $\lim_{m\to\infty} D_m = A$，这就证明了在 $\Omega_{CC}(\mathbf{R})$ 中每个序列都有收敛的子列。

设 $\varepsilon > 0$。存在 $N \in \mathbf{N}$，使对所有的 j，$m > N$ 成立

$$h(D_j,\ D_m) \leqslant \varepsilon$$

因此

$$e(D_m,\ D_j) \leqslant \varepsilon$$

故 $e(A_m,\ D_j) \leqslant \varepsilon$。从而，对所有的 $j > N$，$e(A,\ D_j) \leqslant \varepsilon$。其次，存在 $M \in \mathbf{N}$，使得对所有的 $m > M$ 有 $e(A_m,\ A) \leqslant \varepsilon$。如若不然，则存在序列 (x^i_m)，使对所有的 $i \in \mathbf{N}$，$x^i_m \in A^i_m$，有 $d(x^i_m,\ A) > \varepsilon$，这里 d 为欧几里得度量。由 A^* 的紧性，序列 (x^i_m) 有收敛子列 (y_i)，若 $y_i \to y$，则 $d(y,\ A) \geqslant \varepsilon$，但 $y \in A$，矛盾。从而，存在 $M \in \mathbf{N}$，对所有 $m > M$，有 $e(D_m,\ A) \leqslant \varepsilon$，于是对任意 $m > \max(N,\ M)$ 就有 $h(D_m,\ A) \leqslant \varepsilon$。

相 对 内 部

4.7 定理

注 原文误为 A_1

设C为$\mathbf{R}^n$中凸集，下述情况是等价的：

(a) C是凸代数体。

(b) C是凸体。

(c) $\dim(C)=n$。

证明 (c)$\Longrightarrow$(b)：由定理4.2，C包含一个n-单纯形$S=\mathrm{co}(x_0, x_1, \cdots, x_n)$。我们把证明

$$\mathrm{int}(S)=\left\{\sum_{i=0}^{n}\lambda_i x_i \mid \lambda_i>0(0\leqslant i\leqslant n),\ \sum_{i=0}^{n}\lambda_i=1\right\}$$

留给读者。因此，$\mathrm{int}(S)\neq\phi$。导出$\mathrm{int}(C)\neq\phi$。

(b) $\Longrightarrow$ (a)：由于定理2.23(c)，这是正确的。

(a)$\Longrightarrow$(c)：假设C是一个凸代数体，且$\dim(C)<n$。这蕴涵$\mathrm{aff}(C)\neq\mathbf{R}^n$。设$c\in C^i$及$x\in\mathbf{R}^n\backslash\mathrm{aff}(C)$，令$m$为通过$x$和的直线，就有$m\cap\mathrm{aff}(C)=\{c\}$，因此，$m\cap C=\{c\}$，矛盾。∎

4.8 定义

设E为线性拓扑空间，$A\subset E$。A的相对内部$\mathrm{ri}(A)$定义成视A为$\mathrm{aff}(A)$（具有相对拓扑）的子集时，A的内部。

注 在$\mathbf{R}^n$中引入相对闭包的概念是没有用处的。事实上，因为$\mathbf{R}^n$中每一个仿射集都是闭集，A的闭包$\overline{A}$与A作为$\mathrm{aff}(A)$的子集时的闭包是一样的。

定理。设$C\subset\mathbf{R}^n$是凸的。

(a) 若$C\neq\phi$，则$\mathrm{ri}(C)\neq\phi$且$\dim(\mathrm{ri}(C))=\dim(C)$。

(b) $\overline{\mathrm{ri}(C)}=\overline{C}$且$\mathrm{ri}(\overline{C})=\mathrm{ri}(C)$。

(c) 若$x\in\mathrm{ri}(C)$，且$y\in\overline{C}$，则$[x, y\rangle\subset\mathrm{ri}(C)$。

证明。(a)由§4.7。应用(a)和定理2.27，即得(b)。由定

理2.23推出(c)。

4.9

下面我们研究运算'ri'的某些性质。

定理. 设$C\subset\mathbf{R}^n$是凸的，T：$\mathbf{R}^n\longrightarrow\mathbf{R}^m$为线性映射．则

$$\mathrm{ri}(TC)=T(\mathrm{ri}(C))\text{ 且 }\overline{TC}\supset T(\overline{C}).$$

证明. 从$\mathbf{R}^n$到$\mathbf{R}^m$的任何线性映射是连续的．第二个关系由T的连续性就可得出。

对$\mathrm{ri}(C)$应用第二个关系，由§4.8我们有

$$\overline{T(\mathrm{ri}(C))}\supset T(\overline{\mathrm{ri}(C)})=T(\overline{C})\supset TC\supset T(\mathrm{ri}(C)).$$

由此得$\overline{TC}=\overline{T(\mathrm{ri}(C))}$.再用§4.8，有$\mathrm{ri}(TC)=\mathrm{ri}(\overline{TC})=\mathrm{ri}(\overline{T(\mathrm{ri}(C))})=\mathrm{ri}(T(\mathrm{ri}(C)))$，因此，

$$\mathrm{ri}(TC)\subset T(\mathrm{ri}(C)). \tag{25}$$

设$x\in T(\mathrm{ri}(C))$。存在$x_1\in\mathrm{ri}(C)$，使$x=Tx_1$。由§4.8，$\mathrm{ri}(TC)\neq\phi$，因此存在$y\in\mathrm{ri}(TC)$及$y_1\in C$使$y=Ty_1$。因为$x_1\in\mathrm{ri}(C)$，故有$z_1\in C$，使$x_1\in\langle z_1, y_1\rangle$。设$z=Tz_1$。我们有$z\in TC$及$x\in\langle z, y\rangle$；从§4.8得出$x\in\mathrm{ri}(TC)$。于是

$$T(\mathrm{ri}(C))\subset\mathrm{ri}(TC). \tag{26}$$

结合(25)和(26)式，得出所述结果。

4.10 定理

设C，D为$\mathbf{R}^n$中凸集，$\lambda\in\mathbf{R}$。则：

(a) $\mathrm{ri}(\lambda C)=\lambda\mathrm{ri}(C)$

(b) $\mathrm{ri}(C+D)=\mathrm{ri}(C)+\mathrm{ri}(D)$

进而，若$\mathrm{ri}(C)\cap\mathrm{ri}(D)\neq\phi$，则：

(c) $\overline{C\cap D}=\overline{C}\cap\overline{D}$

(d) $\mathrm{ri}(C\cap D)=\mathrm{ri}(C)\cap\mathrm{ri}(D)$

证明．(a) 应用§4.9到T：$\mathrm{R}^n\to\mathrm{R}^n$，$T$定义为：$Tx=\lambda x$。

(b) 应用§4.9到T：$\mathrm{R}^n\times\mathrm{R}^n\to\mathrm{R}^n$，定义$T(x,\ y)=x+y$。从而推出$\mathrm{ri}(C+D)=\mathrm{ri}(T(C\times D))=T(\mathrm{ri}(C\times D))$。请读者证明

$$\mathrm{aff}(C\times D)=\mathrm{aff}(C)\times\mathrm{aff}(D)$$

及

$$\mathrm{ri}(C\times D)=\mathrm{ri}(C)\times\mathrm{ri}(D)。$$

于是

$$\mathrm{ri}(C+D)=T(\mathrm{ri}(C)\times\mathrm{ri}(D))=\mathrm{ri}(C)+\mathrm{ri}(D)。$$

(c)，(d)。我们有$\overline{C\cap D}=\overline{C}\cap\overline{D}$．设$x\in\mathrm{ri}(C)\cap\mathrm{ri}(D)$，$y\in\overline{C}\cap\overline{D}$。由§4.8，$[x,\ y\rangle\subset\mathrm{ri}(C)\ \cap\mathrm{ri}(D)$，因此$y\in\overline{\mathrm{ri}(C)\cap\mathrm{ri}(D)}$。从而

$$\overline{C}\cap\overline{D}\subset\overline{\mathrm{ri}(C)\cap\mathrm{ri}(D)}\subset\overline{C\cap D}\subset\overline{C}\cap\overline{D}$$

故$\overline{C\cap D}=\overline{C}\cap\overline{D}$，(c)得证．再由§4.8得

$$\mathrm{ri}(\mathrm{ri}(C)\cap\mathrm{ri}(D))=\mathrm{ri}(C\cap D)$$

得

$$\mathrm{ri}(C\cap D)\subset\mathrm{ri}(C)\cap\mathrm{ri}(D)。$$

现设$z\in\mathrm{ri}(C)\cap\mathrm{ri}(D)$，$w\in\mathrm{ri}(C\cap D)$。存在 $u_1\in C$，$u_2\in D$，使$z\in\langle u_1,\ w\rangle\cap\langle u_2,\ w\rangle$，因此存在$u\in C\cap D$，使$z\in\langle u,\ w\rangle$。根据§4.8，$Z\in\mathrm{ri}(C\cap D)$。于是 $\mathrm{ri}(C)\cap\mathrm{ri}(D)\subset\mathrm{ri}(C\cap D)$，(d)得证。

注． 我们留给读者证明（用一个反例）在(c)和(d)中，条件$\mathrm{ri}(C)\cap\mathrm{ri}(D)\neq\phi$不可省略．

R^n 中 的 分 离

4.11

下述定理是分离定理（定理3.8）在有限维线性拓扑空间的一种改进．它给出了在R^n中真分离的一个充分必要条件．

定理（分离定理）．设C，$D\subset\mathrm{R}^n$是凸的和非空的，则在R^n中存在超平面真分离C和D，当且仅当$\mathrm{ri}(C)\cap\mathrm{ri}(D)=\phi$．

证明．我们留给读者证明，当$n=1$时结论成立．现在假定$n\geqslant 2$．设$A=C-D$．A是凸的（见§2.4，性质(a)）且非空．由定理4.10，$\mathrm{ri}(A)=\mathrm{ri}(C)-\mathrm{ri}(D)$，因此有

$$\mathrm{ri}(C)\cap\mathrm{ri}(D)=\phi\Longleftrightarrow 0\bar{\in}\mathrm{ri}(A)$$

充分性：设$B=\mathrm{ri}(A)$．B是相对开集（即在$\mathrm{aff}(B)$中是开的），凸集（见定理2.23）且非空（见§4.8）．有$0\bar{\in}B$．在下述引理中我们将证明必存在超平面$H=f^{-1}(0)$，使$0\in H$，$H\cap B=\phi$．由于B是凸的，不失一般性，我们假设$f(\mathrm{ri}(A))=f(B)>0$，由此得出$f(A)\geqslant 0$．于是

$$(\forall c\in C)\,(\forall d\in D)\qquad f(c)\geqslant f(d)$$

且

$$(\text{存在}c\in C)\,(\text{存在}d\in D)\qquad f(c)>f(d).$$

令$\gamma=\inf\{f(c)\mid c\in C\}$，则$C$和$D$被超平面$f^{-1}(\gamma)$真分离．

必要性：设$H=f^{-1}(\alpha)$是真分离C和D的超平面．不失一

般性，我们假设$f(C)\geqslant\alpha$，$f(D)\leqslant\alpha$ 且对某个 $c\in C$，$f(c)>\alpha$。由此得出$f(A)\geqslant 0$及对某个$a\in A$，$f(a)>0$。设$x\in \mathrm{ri}(A)$。存在$\delta>0$，使得$[x,x+\delta(x-a)]\subset A$，因此$f(x+\delta(x-a))\geqslant 0$，从而$(1+\delta)f(x)\geqslant\delta f(a)$且$f(x)>0$。就有$f(\mathrm{ri}(A))>0$，所以 $0\overline{\in}\mathrm{ri}(A)$。

引理. 设$B\subset \mathbf{R}^n$是凸的和相对开的，且$0\overline{\in}B$。则存在$\mathbf{R}^n$中的一个超平面H，使得$0\in H$，$H\cap B=\phi$。

证明. 我们证明下述命题：若F是$\mathbf{R}^n$的线性子空间，满足$0\leqslant\dim(F)\leqslant n-2$且$F\cap B=\phi$，则存在$\mathbf{R}^n$中包含$F$的超平面$H$，使得$H\cap B=\phi$（令$F=\{0\}$，即得引理）。

首先考虑$n=2$的情形。我们有$F=\{0\}$且$0\overline{\in}B$。我们寻找的超平面是$\mathbf{R}^n$中过原点0但不与B相交的直线。当$\dim(B)=-1,0,1$时，该直线的存在性是平凡的。而在$\dim(B)=2$时，可由定理3.8得出（因为在这种情形，B是$\mathbf{R}^2$中的开子集）。

其次，我们考虑$n>2$的情形。设S是$\mathbf{R}^n$的二维线性子空间，使得$S\perp F$，令$B_1=S\cap(B+F)$；B_1是凸集。由定理4.10，$\mathrm{ri}(B+F)=\mathrm{ri}(B)+\mathrm{ri}(F)=B+F$。因为$\mathrm{ri}(S)=S$，再次用定理4.10，就有$B_1=\phi$或$\mathrm{ri}(B_1)=S\cap(B+F)=B_1$。在两种情形下，$B_1$都是相对开的。因为$0\overline{\in}B_1$，由已证的$n=2$情形，存在$S$中的直线$m$通过0而不与$B_1$相交。设$S_1=F+m$；$S_1$是$\mathbf{R}^n$中包含$F$的线性子空间，且$\dim(S_1)=\dim(F)+1$。假定$S_1\cap B\neq\phi$。则存在$f\in F$，$a\in m$，$b\in B$，使$f+a=b$，即$a=b-f$，发生矛盾（因为$m\cap B_1=\phi$）。因此$S_1\cap B=\phi$。继续用此方法，经有限步后，可得超平面$H$，使得$H\supset F$，$H\cap B=\phi$。

4.12 应用：支撑超平面

设E是线性拓扑空间，$A\subset E$。A 的相对边界 $\mathrm{rb}(A)$ 定义为：视A为 $\mathrm{aff}(A)$（具有相对拓扑）的子集时A的边界。若 $A\subset \mathbf{R}^n$，则$\mathrm{rb}(A)=\overline{A}\setminus \mathrm{ri}(A)$。

下面定理及其证明与§3.10实际上是相同的。

定理. 设$C\subset \mathbf{R}^n$为非空凸集，$x\in \mathrm{rb}(C)$。则 C 在x存在非平凡的支撑超平面。

4.13 应用：极端点

在§2.10中，我们已引进了极端点的概念。下面将证$\mathbf{R}^n$中的紧凸集完全由它的极端点确定。我们从一个引理开始。

引理. 设A是$\mathbf{R}^n$的非空有界子集，A至少含两个点，则$\mathrm{ri}(A)\subset \mathrm{co}(\mathrm{rb}(A))$。

证明. 设$x\in \mathrm{ri}(A)$。存在$y\in A$，$y\neq x$*。设m为通过 x和y的直线，集 $m\cap A$有界且包含一条以x为相对内点的线段。从而有p，$q\in m\cap \mathrm{rb}(A)$，使得$x\in[p, q]$。

定理（**Minkowski 定理**）. 设$C\subset \mathbf{R}^n$为非空紧凸集，E 是C的极端点集。则$E\neq\phi$且$C=\mathrm{co}(E)$。

证明. 设 $d=\dim(C)$。对d采用归纳法证明。

若 $d=0$，C 由一个点组成，定理是平凡的。现设$\dim(C)=k$，且当$d<k$时命题成立。因 C 有界，$\mathrm{rb}(C)\neq\phi$。设 $x_0\in \mathrm{rb}(C)$，$H=f^{-1}(\alpha)$是C在x_0非平凡的支撑超平面（见§4.12）

---

* 译注：原文错为$y\in x$。

使$f(C)\leqslant\alpha$。$H\cap C$为非空紧凸集。读者易证$\dim(H\cap C)<k$。按归纳假定，$H\cap C$的极端点集E_1非空，且$H\cap C=\mathrm{co}(E_1)$。设$e\in E_1$。假定存在x，$y\in C$，$\lambda\in\langle 0,1\rangle$使$e=\lambda x+(1-\lambda)y$。因为$e\in H$，有$f(e)=\alpha$，这样$\lambda f(x)+(1-\lambda)f(y)=\alpha$。由$f(x)\leqslant\alpha$，$f(y)\leqslant\alpha$，得出$f(x)=f(y)=\alpha$，因此$x$，$y\in H\cap C$。由$e\in E_1$，就有$x=y=e$。于是$e\in E$，即$E_1\subset E$（因此$E\neq\phi$）。

由于C是闭的，我们有$x_0\in H\cap C$，从而$x_0\in\mathrm{co}(E_1)\subset\mathrm{co}(E)$。由此$\mathrm{rb}(C)\subset\mathrm{co}(E)$。

由上面引理，我们有$\mathrm{ri}(C)\subset\mathrm{co}(\mathrm{rb}(C))$，因此$\mathrm{ri}(C)\subset\mathrm{co}(E)$。从而$C\subset\mathrm{co}(E)$。便有$C=\mathrm{co}(E)$，即$d=k$时命题成立。

注。 结合Minkow sik 定理与定理4.2，可知每点$x\in C$是C的至多$n+1$个极端点的凸组合。

多 面 锥

4.14

若$x=(\xi_1,\xi_2,\cdots,\xi_n)\in\mathbf{R}^n$，以$x\geqslant 0$表述$(\forall i)\xi_i\geqslant 0$。用$P^n$表示$\mathbf{R}^n$中非负卦限$\{x\in\mathbf{R}^n|\,x\geqslant 0\}$。$\mathbf{R}^n$的闭半空间是$\mathbf{R}^n$中形如$\{x\in\mathbf{R}^n|\,(x|\,p)\leqslant\alpha\}$的子集，这里$p\in\mathbf{R}^n$，$p\neq 0$，$\alpha\in\mathbf{R}$。当$\alpha=0$时，这样的集合称为含边界点0的闭半空间。

定义。 $\mathbf{R}^n$中的多面锥是$\mathbf{R}^n$的有限多个含边界点0的闭半空间之交。

多面锥是包含原点0的凸锥。

设K是$\mathbf{R}^n$中的多面锥，它可表为有限不等式组

$$(x|\,p_i)\leqslant 0\quad(1\leqslant i\leqslant k)$$

的解集。这里$p_i\in\mathbf{R}^n$，$p_i\neq 0$（$1\leqslant i\leqslant k$）。如果定义映射 T：$\mathbf{R}^n\rightarrow\mathbf{R}^k$为

$$(Tx)_i=(x\mid p_i)\qquad(x\in\mathbf{R}^n,\ 1\leqslant i\leqslant k)$$

则有$K=\{x\in\mathbf{R}^n\mid Tx\leqslant 0\}=-T^{-1}p^k$

4.15

下面要用到这样一个事实：$\mathbf{R}^n$中有限个点的所有非负线性组合所成之集是包含0的凸锥。其证明留给读者。

定义。 $\mathbf{R}^n$中的凸锥K称为是有限生成的，如果存在$\mathbf{R}^n$的一个有限子集$\{a_1, a_2, \cdots, a_k\}$，它叫$K$的生成集，使得

$$K=\{\lambda_1a_1+\lambda_2a_2+\cdots+\lambda_ka_k\mid \lambda_1,\ \lambda_2,\ \cdots,\ \lambda_k\geqslant 0\}$$

由此，$\mathbf{R}^n$中的有限生成凸锥也可以定义成下面形式

$$K=\{Tx\mid x\geqslant 0\}=TP^k$$

这里$k\in\mathbf{N}$，T是$\mathbf{R}^k$到$\mathbf{R}^n$的线性映射。

定理。 $\mathbf{R}^n$中的有限生成凸锥的极是一个多面锥。

证明。 设$K=T\mathbf{P}^k$为$\mathbf{R}^n$中有限生成凸锥。对所有$y\in\mathbf{R}^n$，我们有

$$(T\mathbf{P}^k\mid y)\leqslant 0\Longleftrightarrow(\mathbf{P}^k\mid T^ty)\leqslant 0\Longleftrightarrow T^ty\leqslant 0$$

这里T^t：$\mathbf{R}^n\longrightarrow\mathbf{R}^k$为$T$的伴随。从而可得

$$K^0=\{y\in\mathbf{R}^n\mid T^ty\leqslant 0\}.$$

4.16

下面两个定理将使我们能证明多面锥与有限生成凸锥是一致的。

定 理．设T：$\mathbf{R}^n \to \mathbf{R}^k$为线性映射．$b \in \mathbf{R}^k$．$A = \{x \in \mathbf{R}^n | Tx \leqslant b\}$．则：

(a) A至多有有限个极端点．

(b) 若A非空有界，则A是凸多胞形（见§2.8）．

证明．(a) 设$b = (\beta_1, \beta_2, \cdots, \beta_k)$，$A$ 是不等式组

$$(x | p_i) \leqslant \beta_i$$

的解集，这里$p_i \in \mathbf{R}^n$（$1 \leqslant i \leqslant k$）．设$E$为$A$的极端点集．假定$x$，$y \in E$．定义$I(x) = \{i | (x | p_i) < \beta_i\}$，类似地定义$I(y)$．

假设$I(x) = I(y)$．存在$\varepsilon > 0$，使得

$$(x + \varepsilon(x - y) | p_i) \leqslant \beta_i \tag{27}$$

对任何$i \in I(x)$．若$i \overline{\in} I(x)$，我们有$(x | p_i) = (y | p_i) = \beta_i$，因此 $(x + \varepsilon(x - y) | p_i) = \beta_i$．从而对所有的$i$，(27) 式成立．于是$x + \varepsilon(x - y) \in A$．我们有

$$x = \frac{\varepsilon}{1 + \varepsilon} y + \frac{1}{1 + \varepsilon}[x + \varepsilon(x - y)].$$

因为x，$y \in E$，推出$x = y$．这就是说$x \neq y$蕴涵$I(x) \neq I(y)$．由于集 $\{1, 2, \cdots, k\}$仅有有限多个子集，这就证明了所述结果．

(b) A为闭凸集．从定理的假设得出A是非空 紧集．应用Minkowski定理（见§4.13）便得出所述结果．

4.17 定理

设K为$\mathbf{R}^n$中有限生成凸锥．则：

(a) K是有限个有限生成凸锥的并，其中每个具有线性无关生成集．

(b) K是闭的。

证明. (a)(见 Caratheodory定理证明)。存在a_1，a_2，…$a_p \in \mathbf{R}^n$，使

$$K = \{\lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_p a_p | \lambda_1, \lambda_2, \cdots, \lambda_p \geqq 0\}$$

设$x \in K$，$x = \lambda_1 a_1 + \lambda_2 a_2 + \cdots + \lambda_p a_p$，这里 λ_1，λ_2，…，$\lambda_p \geqq 0$。若a_1，a_2，…，a_p是线性相关的，则存在不全为0的 μ_1，μ_2，…，$\mu_p \in \mathbf{R}$，使

$$\mu_1 a_1 + \mu_2 a_2 + \cdots + \mu_p a_p = 0$$

对所有$\rho \in \mathbf{R}$，我们有

$$x = \sum_{i=1}^{p} (\lambda_i - \rho\mu_i) a_i.$$

集 $\{\sigma \in \mathbf{R} | \sigma\mu_i \leqslant \lambda_i (1 \leqslant i \leqslant p)\}$是形如$\langle -\infty, \alpha]$，$[\alpha, \beta]$（这里允许$\alpha = \beta = 0$）或者$[\alpha, +\infty\rangle$的区间。令$\rho = \alpha$，我们得到$x$作为$K$的至多$p-1$个生成元的非负线性组合的表达式。重复这个过程，经有限步后，就得知x为K的线性无关生成元的非负线性组合。这证明了所述结果，因为K的生成集仅有有限多个子集。

(b) 若a_1，a_2，…，a_p 线性无关，则锥

$$\{\lambda_1 a_1 + \lambda_2 a_2 + \cdots \lambda_p a_p | \lambda_1, \lambda_2, \cdots, \lambda_p \geqslant 0\}$$

是$\mathbf{R}^n$中闭子集，因为它与闭集

$$\{(\xi_1, \xi_2, \cdots, \xi_n) \in \mathbf{R}^n | \xi_1, \xi_2, \cdots, \zeta_p \geqslant 0\}$$

同胚。由(a)知K是闭的。

4.18 定理

设$K \subset \mathbf{R}^n$。则K是多面锥当且仅当K是有限生成凸锥。

证明。必要性： 设 $K=\{x\in\mathbf{R}^n \mid Tx\leqslant 0\}$（这里 $T:\mathbf{R}^n\to\mathbf{R}^k$为线性映射）是一个多面锥，我们有

$$K=T^{-1}(T\mathbf{R}^n\cap(-P^k))。$$

设$A=T\mathbf{R}^n\cap(-P^k)$，$B=\{(\eta_1,\eta_2,\cdots,\eta_k)\in A \mid \sum_{i=1}^{k}\eta_i\geqslant -1\}$。$B$非空（因为 $0\in B$）且

$$A=\mathbf{R}_+B \tag{29}$$

这里$\mathbf{R}_+$是所有非负实数组成的集。B是不等式组

$$\begin{cases}(y\mid p_i)\leqslant 0 & (1\leqslant i\leqslant s)\\ (y\mid -p_i)\leqslant 0 & (1\leqslant i\leqslant s)\\ (y\mid e_i)\leqslant 0 & (1\leqslant i\leqslant k)\\ (y\mid -e)\leqslant 1 \end{cases}$$

的解集。这里p_1，p_2，…，p_s是$\mathbf{R}^k$中$T\mathbf{R}^n$ 的正交补的基。e_1，e_2，…，e_k是$\mathbf{R}^k$中的单位向量($e_1=(1,0,0,\cdots,0)$，等等)$e=(1,1,\cdots,1)$。读者易证B有界。由§4.16，存在a_1，a_2，…，$a_p\in\mathbf{R}^k$ 使得

$$B=\mathrm{co}(a_1,a_2,\cdots,a_p)。$$

利用（29）式，我们得出A是由a_1，a_2，…，a_p生成的有限生成凸锥。存在b_1，b_2，…，$b_p\in\mathbf{R}^n$，使$Tb_i=a_i(1\leqslant i\leqslant p)$。若 $x\in K$，存在λ_1，λ_2，…，$\lambda_p\geqslant 0$有

$$Tx=\sum_{i=1}^{p}\lambda_i a_i=\sum_{i=1}^{p}\lambda_i Tb_i$$

因此

$$x-\sum_{i=1}^{p}\lambda_i b_i\in T^{-1}\{0\}。$$

从而K是由b_1，b_2，…，b_p，c_1，c_2，…，c_q，$-c_1$，$-c_2$，…，

$-c_q$生成的有限生成凸锥，这里c_1，c_2，…，c_q是T的核$T^{-1}\{0\}$的基。

充分性：设K为有限生成凸锥，由定理4.17，K是闭的。从§3.13得

$$K = K^{00} = (K^0)^0.$$

按照§4.15，K^0 是多面锥，因此按前所证，K^0是有限生成凸锥。最后，再用§4.15，K^{00}是多面锥。

注。 从上面已证及§4.15，我们得出集$\{Tx \mid x \geqslant 0\}$和$\{y \mid T^t y \leqslant 0\}$之一是另一个的极。

4.19 应用：Farkas引理

设A：$\mathbf{R}^n \to \mathbf{R}^h$为线性映射，令$K = \{x \mid A^t x \leqslant 0\}$。

我们有

$$b \in K^0 \Longleftrightarrow (\forall x \in \mathbf{R}^h)(A^t x \leqslant 0 \Longrightarrow (b \mid x) \leqslant 0)$$

从§4.18的注推出$K^0 = \{Ay \mid y \geqslant 0\}$，因此

$$(\text{存在} y \geqslant 0)\, b = Ay \Longleftrightarrow (\forall x \in \mathbf{R}^h)(A^t x \leqslant 0 \Longrightarrow (b \mid x) \leqslant 0). \tag{30}$$

（30）式称为$Farkas$引理。此引理也可表述成择一定理：

对每个线性映射A：$\mathbf{R}^n \to \mathbf{R}^h$及每个$b \in \mathbf{R}^h$，或者

（Ⅰ） $A^t x \leqslant 0$，$(b \mid x) > 0$有解$x \in \mathbf{R}^h$，或者

（Ⅱ） $b = Ay$，$y \geqslant 0$有解$y \in \mathbf{R}^n$

但两者决不同时成立。

练　　习

V是R上的线性空间。

1　设$A\subset V$，$\dim(A)=k$，$a\in A$。证明每个$x\in \mathrm{co}(A)$含于某个顶点在A中的k-单纯形内，a亦是顶点之一。

2　设$A\subset V$，证明$\mathrm{co}(A)$是顶点属于A的所有有限维单纯形之并。

3　设C_1，C_2，…，C_n是V中不相交的凸子集，$x\in \mathrm{co}\left(\bigcup_{i=1}^{n} C_i\right)$。证明$x$含于一个至多有一个顶点在$C_i$的单纯形中。

4　设Γ是$\mathbf{R}^n$中所有非空紧子集的集族，具有Hausdorff距离。(A_i)是Γ中序列，$A\in\Gamma$，使得$A_i\to A(i\to\infty)$。设(x_i)是$\mathbf{R}^n$中的序列，$x_i\in A_i(i\in\mathbf{N})$。假定$x\in\mathbf{R}^n$使得$x_i\to x(i\to\infty)$证明$x\in A$。

5　设Γ定义如练习4．(A_i)是Γ中一序列，$A\in\Gamma$，使得$A_i\to A$ $(i\to\infty)$．如果每个A_i是凸的，证明A是凸的。

6　(a) 设$C\subset\mathbf{R}^n$是凸的。证明．

$C^a=\overline{C}$．

(b) 举出一个线性拓扑空间中的凸子集C，使得$C^a\neq\overline{C}$．

7　设$C\subset\mathbf{R}^n$是凸集。证明C是闭的当且仅当对$\mathbf{R}^n$中每条直线m，$C\cap m$是闭集。

8　设$C\subset\mathbf{R}^n$是凸集，$A\subset\mathbf{R}^n$是开集．证明$A\cap\overline{C}\neq\phi$当且仅当$A\cap \mathrm{ri}(C)\neq\phi$。

9　设C，$D\subset\mathbf{R}^n$是凸集，$C\subset\overline{D}$，$C\cap \mathrm{ri}(D)\neq\phi$。证明$\mathrm{ri}(C)\subset \mathrm{ri}(D)$。

10 (a) 证明定理4.10 (c) 可推广到任意凸集族。

(b) 证明定理4.10(d)可推广到任意多于两个凸集的有限集族的情形，但对无限个凸集所成的集族并不成立。

11 证明Radon定理：设$A\subset R^n$至少包含$n+2$个点，则存在A_1，$A_2\subset R^n$，使得$A_1\cap A_2=\phi$，$A_1\cup A_2=A$，$co(A_1)\cap co(A_2)\neq\phi$。

12 设$K\subset R^n$为非空凸锥。C是R^n的非空凸子集，满足$C\cap K=\phi$。证明存在$y\in K^0$，$y\neq0$，使得对任何$c\in C$，$(c\ y)\geqslant0$。

13 设K_1，$K_2\subset R^n$为多面锥。证明

$$(K_1\cap K_2)^0=K_1^0+K_2^0$$

（见第三章练习6）。

14 应用分离定理到$\{b\}$和AP^n，证明Farkas引理。

（提示：$b\ \overline{\in}AP^n\Longleftrightarrow$存在超平面严格分离$\{b\}$和$AP^n$）。

15 证明Gordan引理：对每个线性映射A：$R^n\to R^k$，或者

（Ⅰ）存在$x\in R^k$，使 $A^tx>0$；或者

（Ⅱ）存在$y\in R^n$，使得

$$Ay=0,\ y\geqslant0,\ y\neq0;$$

但两者决不同时成立。

注　　释

1 凸集几何学较凸函数分析出现得更早。其早期结果见T.Bonnesen and W. Fenchel, *Theorie der konvexen Körper*, re-issued by Chelsea, New York,1971。

在此领域中其它有价值的参考书是

H. G. Eggleston, *Convexity*, Cambridge, Cambridge University Press, 1969,

F. A. Valentine, *Conex Sets*, New York, McGraw-Hill, 1964,

K. Leichtweisz, *Konvexe Mengen*, Berlin, Springer-Verlag, 1980。

一般说来，我们仅给出在凸分析中有用的那些与凸集有关的结果。

2 对无穷维空间也有§4.3的类似定理。这就是Mazur定理：若E为Banach空间，$A\subset E$是紧集，则$\overline{\text{co}}\,(A)$亦是紧集。

3 R^r中紧凸子集的体积和表面积是$\Omega_{cc}(\mathrm{R})$上的连续函数（见§4.6）。从Blaschke收敛定理得出$\Omega_{cc}(\mathrm{R})$上任何实连续函数存在极小。这个事实可用于研究等周问题：找出表面积一定而包含最大体积的集合。

4 Minkowski定理（见§4.13）在无穷维空间中有一个推广。这就是Krein-Milman定理：局部凸空间中的非空紧凸集是它的极端点集的闭凸包。

5 线性空间中的凸性概念已以多种方式得到推广。见 V. W. Bryant and R. J. Webster, Generalizations of the theorems of Radon, Helly, and Caratheodory, *Monatshefte für Mathematik* 73 (1969) 309—15

凸性空间定义为序对$(X,\cdot)$，这里X为非空集，$\cdot$是从$X\times X$到X的所有子集族的映射（显然，$a\cdot b$是线段$[a,\ b]$的内部$\langle a,\ b\rangle$的推广）满足：

(a) $a \cdot b \neq \phi$

(b) $a \cdot b = b \cdot a$

(c) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$

(d) 设$a/b = \{x \in X | a \in b \cdot x\}$。则有：

$$(a/b) \cap (c/d) \neq \phi \Longrightarrow (a \cdot d) \cap (b \cdot c) \neq \phi$$

(e) $a \cdot a = \{a\} = a/a$

(f) $(a \cdot b) \cap (a \cdot c) \neq \phi \Longrightarrow b = c$或$b \in a \cdot c$或$c \in a \cdot b$。

在这样的凸性空间中，能定义独立集和维数的概念，就可推广Caratheodory 定理（见定理4.2），Helly 定理（见定理4.4）和Radon定理（见练习[illegible]）．

基于这些概念的更多的几何理论可在下述文献中找到

W. Prenowitz and J. Jantosciak, *Join Geometries, a Theory of Convex Sets and Linear Geometry*, Berlin, Springer, 1979.

6 在文D. C. Kay和E.W. Womble, Axiomatic Convexity theory and relationships between the Caratheodory, Helly and Radon numbers, *Pacific J. Math.* 38 (1971) 471—85中，凸性空间定义为二元对（X，$\mathbf{B}$），这里X是一个集合，$\mathbf{B}$是X的子集族，满足下列条件：

(a) $\phi \in \mathbf{B}$且$X \in \mathbf{B}$。

(b) $\mathbf{B}$的每个子族的交属于$\mathbf{B}$。

集$A \subset X$的凸包，用$C(A)$表示，定义成

$$C(A) = \bigcap \{ B | B \in \mathbf{B},\ B \supset A \}$$

在这种凸性空间中，定义三种数：

（X，$\mathbf{B}$）称为有Caratheodory数c，如果c是这样的非负整数，对所有的$A \subset X$，

$$C(A)=\bigcup\{C(B)\mid B\subset A,\ |B|\leqslant c\}$$

这里$|B|$是B的基数。

$(X,\ \mathbf{B})$称为有Helly数h，若h是这样的最小非负整数，只要子族的每h个集合有非空的交，$\mathbf{B}$的有限子族亦有非空的交。

$(X,\ \mathbf{B})$称为有Radon数r，如果r是这样的最小正整数，对于任何的$A\subset X$，$|A|\geqslant r$，A可被划分成两个非空子集A_1和A_2，使得$C(A_1)\cap C(A_2)\neq\phi$。

若$X=\mathbf{R}^n$，$\mathbf{B}$是$\mathbf{R}^n$中所有凸子集组成的集族，则$c=h=n+1$且$r=n+2$。

可以证明，在c，h和r存在的凸性空间中，我们有$h+1\leqslant r\leqslant ch+1$。

关于对凸性理论和有关课题建立的更多的公理可在下述文献中找到：

R. E. Jamison, A general theory of converity, PhD Thesis, University of Washington, Seattle, 1974.

G. Sierksma, Relationships between Caratheodory, Helly, Radon and exchange numbers of convexity spaces, *Nieuw Archief voor Wiskunde*(3), XXV(1977)115—132

H. Van Maaren and H. J. P. De Smet, Extremal points, separation and Caratheodory, Helly and Radon numbers in non-real linear spaces, *Proc. Kon. Ned. Akad. Wet.* 84 (= *Indag. Math.* 43)(1981) 207—18.

7　我们定义从集合X到集合Y的多值函数为从X到Y的一个关系，或更确切地说，是从X到Y的幂集$P(Y)$的函数。从X到Y的多值函数的性质，有时可用赋予适当结构的$P(Y)$

(或它的部分)来描述。这里的结构与已给Y的结构相容。这些结构的一个例子是由Hausdorff度量导出的Ω_{cc}上的拓扑(见§4.6)。对多值函数很值得一读的入门文献是

R. E. Smithson, Multifunctions, *Nieuw Archief voor Wiskunde* (3), XX (1972) 31—53.

B. L. McAllister, Hyperspaces and Multifunctions, the first half century (1900—1950), *Nieuw Archief voor Wiskunde*(3), XXVI (1978) 309—29.

一个与Blaschke收敛定理密切相关的定理如下:

设X为完备度量空间,$P(X)$为X的所有非空有界闭子集组成的集族,则具有Hausdorff距离的$P(X)$是一个完备度量空间。

设Y为线性空间,多值函数$f: X\to Y$称为是凸的,如果对每个$x\in X$,$f(x)$是Y的凸子集。关于凸多值函数的理论可在下面文献中找到:

C. Castaing and M. Valadier, *Convex Analysis and Measureable Multifunctions* (*Lecture Notes in Mathematics no.* 580), Berlin, Springer, 1977.

第五章　线性空间上的凸函数

本章，用V表示R上的线性空间，E表示R上具有 Hausdorff 拓扑的线性拓扑空间，V与E中均不止含一个点。

上　图　象

5.1　定义

设X为一集合，$f: X\to\overline{\mathrm{R}}$的上图象 epi($f$)是集合

$$\{(x,\ \lambda)\in X\times \mathrm{R} \mid f(x)\leqslant\lambda\}\ .$$

见图10。

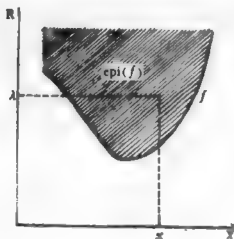


图10

下面，我们常常根据epi(f)的特性来描述 f 的性质。

若 X 为拓扑空间，我们赋 $X\times\mathbf{R}$ 以积拓扑。epi(f)的闭性原来与 f 的下半连续性是相当的（见定理5.3）。

用熟知的方式使 $V\times\mathbf{R}$ 成为一个线性空间，记成 $V\oplus\mathbf{R}$。epi(f)的凸性原来与 f 的凸性是相当的（见定理5.10）。

赋予积拓扑的 $E\oplus\mathbf{R}$ 是一个线性拓扑空间。特别地，若 E 为线性赋范空间（具有范数 $x\to\|x\|$），则 $E\oplus\mathbf{R}$ 上的拓扑是（等价）范数 $(x,\lambda)\to\|x\|+|\lambda|$ 和 $(x,\lambda)\to\max(\|x\|,|\lambda|)$ 之一的范数拓扑。

下 半 连 续 性

设 X 为拓扑空间。

5.2 定义

设 $f: X\to\overline{\mathbf{R}}$，$a\in X$。$f$ 称为在 a 下半连续，如果对每个 $K\in\mathbf{R}$、$K<f(a)$，存在 a 的一个邻域 U，使得 $f(U)>K$。f 称为是下半连续的，若 f 在 X 的每一点均下半连续。

注.

(a) 连续函数是下半连续的。

(b) 设 $a\in X$ 是 X 的聚点，$f(a)=+\infty$，如果 f 在 a 下半连续，则

$$\lim_{x\to a} f(x)=+\infty$$

(c) 若 $f(a)=-\infty$，则 f 在 a 是下半连续的。

5.3 定理

设 $f: X\to\overline{\mathbf{R}}$。下述条件等价

(a) f 是下半连续的。

(b) 对每个 $\lambda\in\mathbf{R}$，$\{x\in X \mid f(x)>\lambda\}$ 是开集。

(c) 对每个 $\lambda\in\mathbf{R}$，$\{x\in X \mid f(x)\leqslant\lambda\}$ 是闭集。

(d) $\operatorname{epi}(f)$ 是闭的（作为 $X\times\mathbf{R}$ 的子集）。

证明 (a) $\Longleftrightarrow$ (b) 是定义 5.2 的直接结果；(b) $\Longleftrightarrow$ (c) 是平凡的。(a) $\Longleftrightarrow$ (d)：定义 F：$X\times\mathbf{R}\to\overline{\mathbf{R}}$ 为 $F(x,\lambda)=f(x)-\lambda$。读者易证 F 是下半连续当且仅当 f 是下半连续。由 (c)，F 的下半连续性也可用对每个 $\mu\in\mathbf{R}$，集

$$\{(x,\lambda) \mid F(x,\lambda)\leqslant\mu\}$$

是闭集来描述。这便证明了所述结果，因为

$$\{(x,\lambda) \mid F(x,\lambda)\leqslant\mu\}=\{(x,\lambda) \mid (x,\lambda+\mu)\in\operatorname{epi}(f)\}$$
$$=\operatorname{epi}(f)-(0,\mu).$$

5.4

下面简单性质的证明留给读者。

(a) 任意的下半连续函数集族的点态上确界是下半连续的。

(b) 若 X 为紧集且 f：$X\to\mathbf{R}\cup\{+\infty\}$ 是下半连续的，则 f 必有极小值（它可以是 $+\infty$）。

(c) 若 f，g 是两个下半连续函数 $X\to\mathbf{R}\cup\{+\infty\}$，$\lambda>0$，则 λf 和 $f+g$ 亦是下半连续函数。

5.5

函数 f：$X\to\overline{R}$ 之上图象的闭包 $\overline{\operatorname{epi}(f)}$ 也是一个上图象。事实上，设 $(x,\lambda)\in\overline{\operatorname{epi}(f)}$，$\mu>\lambda$，假定 $U\times W$ 是 (x,μ) 在

$X\times\mathbf{R}$ 中的一个邻域.存在开区间 $I\subset\mathbf{R}$，使 $\lambda\in I$，$\mu\in I$.因为 $(x,\lambda)\in\overline{\mathrm{epi}(f)}$，存在点 $(y,\sigma)\in\mathrm{epi}(f)$，使 $(y,\sigma)\in U\times I$.我们有 $f(y)\leqslant\sigma<\mu$，故 $(y,\mu)\in\mathrm{epi}(f)\cap(U\times W)$。从而 $(x,\mu)\in\overline{\mathrm{epi}(f)}$。我们得出 $\overline{\mathrm{epi}(f)}$ 和直线 $\{x\}\times\mathbf{R}$ 的交是 ϕ，$[a,+\infty)$，或 $\mathbf{R}$。分别令 $g(x)=+\infty$，$g(x)=\lambda$，和 $g(x)=-\infty$. 则定义了一个上图象是 $\overline{\mathrm{epi}(f)}$ 的函数 g.

定义。函数 f：$X\rightarrow\overline{\mathbf{R}}$ 的下半连续包 $\bar{f}$ 是上图象为 $\overline{\mathrm{epi}(f)}$ 的函数 ($X\rightarrow\overline{\mathbf{R}}$) 且

$$\mathrm{epi}(\bar{f})=\overline{\mathrm{epi}(f)}.$$

5.6

g：$X\rightarrow\overline{\mathbf{R}}$ 称为 f：$X\rightarrow\overline{\mathbf{R}}$ 的弱函数，如果对所有的 $x\in X$，$g(x)\leqslant f(x)$，这等价于 $\mathrm{epi}(g)\supset\mathrm{epi}(f)$。由定理 5.3，$f$ 的每个下半连续弱函数的上图象是闭的（且包含 $\mathrm{epi}(f)$）。可知 $\bar{f}$ 是 f 的最大下半连续弱函数。注意，$\bar{f}$ 也可定义为 f 的所有下半连续弱函数的点态上确界（常值函数 $-\infty$ 是这些弱函数之一）。

5.7

我们也可利用下极限概念引入下半连续性，定义成

$$\liminf_{x\to a} f(x):=\sup_{U}\{\inf f(x)\mid x\in U\setminus\{a\}\}\quad(31)$$

这里 U 属于 a 的邻域系。若 X 为一线性赋范空间，则有

$$\liminf_{x\to a} f(x)=\lim_{\varepsilon\downarrow 0}\inf\{f(x)\mid 0<\|x-a\|<\varepsilon\}.$$

5.8 定理

设$f: X \to \overline{\mathbf{R}}$，$a \in X$。则

(a) f是下半连续的$\Longleftrightarrow f = \bar{f}$；

(b) f在a下半连续$\Longleftrightarrow \liminf_{x \to a} f(x) \geqslant f(a)$；

(c) $\bar{f}(a) = \min\{f(a), \liminf_{x \to a} f(x)\}$；

(d) f在a下半连续$\Longleftrightarrow \bar{f}(a) = f(a)$。

证明. (a) 由定理5.3，我们有：

$$f\text{是下半连续的} \Longleftrightarrow \mathrm{epi}(f) = \overline{\mathrm{epi}(f)}$$

$$\Longleftrightarrow \mathrm{epi}(f) = \mathrm{epi}(\bar{f}) \Longleftrightarrow f = \bar{f}$$

(b) 是定义5.2的直接结果。

(c) 鉴于 (b) 与$\bar{f}$的下半连续性，我们有

$$\bar{f}(a) \leqslant \liminf_{x \to a} \bar{f}(x) \leqslant \liminf_{x \to a} f(x).$$

还有$\bar{f}(a) \leqslant f(a)$，因此$\bar{f}(a) \leqslant \mu$。这里

$$\mu: = \min\{f(a), \liminf_{x \to a} f(x)\}.$$

若设$\bar{f}(a) < \mu$，这蕴涵存在$\lambda \in \mathbf{R}$，满足$\mu > \lambda$，$\bar{f}(a) \in I: = [-\infty, \lambda)$及存在$a$的邻域$U$使得$\inf\{f(x) \mid x \in U \setminus \{a\}\} > \lambda$，故对任意$x \in U$，$f(x) > \lambda$。即$U \times I$为$(a, \bar{f}(a))$（在$X \times \mathbf{R}$中）的一个邻域，且不与$\mathrm{epi}(f)$相交。这与$(a, \bar{f}(a)) \in \mathrm{epi}(\bar{f}) = \overline{\mathrm{epi}(f)}$矛盾。从而$\bar{f}(a) = \mu$。

(d) 是 (b) 和 (c) 的直接结果。

注。 有一种替代的下极限定义（见31）如下：

$$\liminf_{x \to a} f(x): = \sup_{U} \{\inf f(x) \mid x \in U\}.$$

用这种定义，有 $\bar{f}(a)=\liminf_{x\to a} f(x)$．

凸　　性

5.9 定义

函数 f：$V\to\overline{\mathrm{R}}$ 称为凸的，如对所有的 $x, y\in V$，和 λ，μ，$\nu\in\mathrm{R}$，$f(x)<\mu$，$f(y)<\nu$，$0<\lambda<1$，成立

$$f(\lambda x+(1-\lambda)y)<\lambda\mu+(1-\lambda)\nu, \tag{32}$$

见§1.20的定义1.19和第一章练习9．

从定义得到

f 是凸的 $\Longleftrightarrow$ 对 V 中每条直线 m，f 在 m 上的限制是凸的．

设 m 是过 x 和 y 的直线(因此 $m=\{(1-t)x+ty \mid t\in\mathrm{R}\}$)．用 $f\mid m$ 表示 f 在 m 上的限制．定义 f_m：$\mathrm{R}\to\overline{\mathrm{R}}$ 为 $f_m(t)=f((1-t)x+ty)$，则

$f\mid m$ 是凸的 $\Longleftrightarrow f_m$ 是凸的．

5.10 定理

设 f：$V\to\overline{\mathrm{R}}$．下述条件是等价的：

(a)　f 是凸的．

(b)　$\mathrm{epi}(f)$ 是凸的．

(c)　$\{(x, \lambda)\in V\oplus\mathrm{R} \mid f(x)<\lambda\}$ 是凸的．

证明．(a) $\Longleftrightarrow$ (c)：设

$$A:=\{(x,\lambda)\in V\oplus\mathrm{R} \mid f(x)<\lambda\}.$$

由

$$\lambda(x,\mu)+(1-\lambda)(y,\nu)=(\lambda x+(1-\lambda)y,\lambda\mu+(1-\lambda)\nu),$$

有

$$A\text{是凸的}\Longleftrightarrow f(\lambda x+(1-\lambda)y)<\lambda\mu+(1-\lambda)\nu$$

$$\text{对任意的 } f(x)<\mu,\ f(y)<\nu,\ 0<\lambda<1.$$

应用定义 5.9，(a) 与(c)等价得证。

(b)$\Longleftrightarrow$(c)：假设 epi(f) 是凸的．$(x,\mu)\in A$，$(y,\nu)\in A$，$0<\lambda<1$．因为 $f(x)<\mu$，$f(y)<\nu$，我们可选 μ_0，$\nu_0\in\mathbf{R}$，使 $f(x)\leqslant\mu_0<\mu$，$f(y)\leqslant\nu_0<\nu$．于是，$(x,\mu_0)\in$ epi(f)，$(y,\nu_0)\in$ epi(f)，因此

$$(\lambda x+(1-\lambda)y,\ \lambda\mu_0+(1-\lambda)\nu_0)\in \mathrm{epi}(f)$$

故

$$f(\lambda x+(1-\lambda)y)\leqslant\lambda\mu_0+(1-\lambda)\nu_0<\lambda\mu+(1-\lambda)\nu .$$

故 $(\lambda x+(1-\lambda)y,\ \lambda\mu+(1-\lambda)\nu)\in A$，$A$是凸的。

反之，设A是凸的．若 $f(x)\leqslant\mu$，$f(y)\leqslant\nu$，$0<\lambda<1$，则对任意 $\varepsilon>0$，

$$f(x)<\mu+\varepsilon,\qquad f(y)<\nu+\varepsilon$$

因此

$$f(\lambda x+(1-\lambda)y)<\lambda(\mu+\varepsilon)+(1-\lambda)(\nu+\varepsilon)$$
$$=\lambda\mu+(1-\lambda)\nu+\varepsilon.$$

必 $f(\lambda x+(1-\lambda)y)\leqslant\lambda\mu+(1-\lambda)\nu$，epi$(f)$ 是凸的。

5.11 定义

(a) 凸函数 f：$V\to\overline{\mathbf{R}}$的有效域是集 $\{x\in V\mid f(x)<+\infty\}$，用 dom$(f)$表示。

(b) V上的正常凸函数是不恒等于 $+\infty$的凸函数 $V\to\mathbf{R}\cup$

$, +\infty\}$。

(c) V上的非正常凸函数是V上的一个凸函数，且不是正常的。

读者易证凸函数的有效域是凸的。V上的正常凸函数f可以看作是由某个有效域为 $\mathrm{dom}(f)$ 的有限凸函数扩充到整个V上而得（$\mathrm{dom}(f)$ 以外，指定取值$+\infty$；见§1.22）。

5.12

如函数$\mathbf{R}\to\overline{\mathbf{R}}$的情形一样（见§1.23），$E$上的非正常凸函数类亦不难描述。

定理. 设f：$E\to\overline{\mathbf{R}}$为非正常凸函数，则

(a) 对所有$x\in\mathrm{ri}(\mathrm{dom}(f))$，$f(x)=-\infty$。

(b) 若f是下半连续的，则$\mathrm{dom}(f)$是闭的，且$f\mid\mathrm{dom}(f)=-\infty$。

证明. (a) 若$f=+\infty$，命题显然成立。若$f\neq+\infty$，则存在$a\in\mathrm{dom}(f)$，使$f(a)=-\infty$。设$y\in\mathrm{ri}(\mathrm{dom}(f))$，$y\neq a$，令$m$为过$a$和$y$的直线。$f\mid m$是非正常凸函数，且$y\in\mathrm{int}(\mathrm{dom}(f\mid m))$。应用§1.23得到所述结果。

(b) 设$f\neq+\infty$，且$f(a)=-\infty$。假定存在$b\in E$，使$f(b)\in\mathbf{R}$。令m为过a和b的直线。我们有$\langle a,b\rangle\subset\mathrm{int}(\mathrm{dom}(f\mid m))$，因此，由§1.23，对所有的$x\in\langle a,b\rangle$，$f(x)=-\infty$。但鉴于$f$在$b$的下半连续性，有$b$的邻域$U$，使$f(U)>-\infty$，矛盾。从而$f\mid\mathrm{dom}(f)=-\infty$。

因为$\mathrm{dom}(f)=\{x\in E\mid f(x)\leqslant 0\}$，故$\mathrm{dom}(f)$的闭性由定理5.3即得。

5.13 定义

函数f：$V\to \mathbf{R}$ 称为是严格凸的，如果对任意的x，$y\in V$及所有的$\lambda\in\langle 0, 1\rangle$

$$f(\lambda x+(1-\lambda)y)<\lambda f(x)+(1-\lambda)f(y).$$

5.14

下述简单性质的证明留给读者。

(a) 如果f和g是V上的正常凸函数，$\lambda>0$，$\mu>0$，则$\lambda f+\mu g$是凸的。

(b) V上有限多个正常凸函数的和是凸函数。

(c) V上凸函数序列（点态）收敛的极限是凸函数。

(d) 设f是V上的正常凸函数，则

$$f(\sum_{i=1}^{n}\lambda_i x_i)\leqslant\sum_{i=1}^{n}\lambda_i f(x_i)$$

其中$x_i\in V$，$\lambda_i\geqslant 0$，$(1\leqslant i\leqslant n)$，$\sum_{i=1}^{n}\lambda_i=1$。

(e) 任一V上凸函数集族的点态上确界是凸的。

5.15 例：指标函数

设$A\subset V$，A的指标函数δ_A：$V\to\overline{\mathbf{R}}$ 定义为

$$\delta_A(x)=\begin{cases}0 & \text{如果 } x\in A\\ +\infty & \text{如果 } x\in V\backslash A.\end{cases}$$

我们把下述定理的证明留给读者。

定理。设$A\subset E$。则：

(a) A是凸的$\Longleftrightarrow\delta_A$是凸的.

(b) A是闭的$\Longleftrightarrow\delta_A$是下半连续的.

5.16 凸包

设$A\subset V\times\mathbf{R}$是凸的。定义函数$f$：$V\to\overline{\mathbf{R}}$为

$$f(x)=\inf\{\lambda \mid (x,\lambda)\in A\} \tag{33}$$

（这里，$\inf\phi=+\infty$）。读者易证f是凸的；且f是V上其上图象包含A的最大凸函数。

现设$A=\mathrm{co}\,(\mathrm{epi}(g))$，$g$是任意函数$V\to\overline{\mathbf{R}}$。这时，由（33）定义的函数称为$g$的凸包$\mathrm{co}(g)$。即

$$\mathrm{co}(g)(x)=\inf\{\lambda \mid (x,\lambda)\in \mathrm{co}(\mathrm{epi}(g))\}$$

（见图11）。$\mathrm{co}(g)$是g的最大弱凸函数。如果g是函数$V\to\mathbf{R}\cup\{+\infty\}$，$\mathrm{co}(g)$也可定义成：

$$\mathrm{co}(g)(x)=\inf\left\{\sum_{i=1}^{n}\lambda_i g(x_i)\right\} \qquad (x\in V)$$

这里下确界是关于x作为V中点的有限凸组合$\sum_{i=1}^{n}\lambda_i x_i$的所有表示式来取的（证明留给读者）.

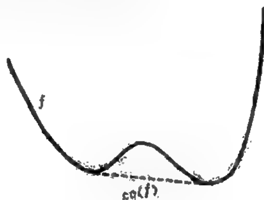


图11

设f：$E\to\mathbf{R}$是凸的。epi(f)的凸性蕴涵着$\overline{\text{epi}(f)}$的凸性。因为epi($\bar{f}$)$=\overline{\text{epi}(f)}$，从而$\bar{f}$是凸的。

可以得出，若g是任意函数$E\to\overline{\mathbf{R}}$，则$\overline{\text{co}(g)}$是$g$的最大下半连续弱凸函数。

5.17 下确界卷积

设f，g是V到$\overline{\mathbf{R}}$的凸函数。集合epi (f) + epi (g) 是凸的。定义f和g的下确界卷积$f\Box g$为

$$(f\Box g)(x)=\inf\{\lambda \mid (x,\lambda)\in \text{epi}(f)+\text{epi}(g)\}\quad (x\in V).$$

$f\Box g$是凸函数$V\to\overline{\mathbf{R}}$。若$f$，$g$是$V$到$\mathbf{R}\cup\{+\infty\}$的函数，定义$f\Box g$为

$$(f\Box g)(x)=\inf\{f(x_1)+g(x_2)\mid x_1+x_2=x\}.$$

或

$$(f\Box g)(x)=\inf\{f(y)+g(x-y)\mid y\in V\}.$$

这与积分卷积公式类似。本段所用术语正是由此引生。

5.18 例

(a) 设B：$V\times V\to\mathbf{R}$为双线性函数，即对所有x，y，$z\in V$及所有的α，$\beta\in\mathbf{R}$，有

$$\begin{cases}B(\alpha x+\beta y,\ z)=\alpha B(x,z)+\beta B(y,z)\\ B(x,\ \alpha y+\beta z)=\alpha B(x,y)+\beta B(x,z)\end{cases}$$

成立。如果对每个$x\in V$，$B(x,x)\geqslant 0$，则函数$x\to B(x,x)$是凸的。事实上，对所有x，$y\in V$及所有的$\lambda\in\mathbf{R}$

$$B(\lambda x+(1-\lambda)y,\ \lambda x+(1-\lambda)y)=\lambda^2B(x,x)$$
$$+\lambda(1-\lambda)\{B(x,y)+B(y,x)\}+(1-\lambda)^2B(y,\ y).$$

我们有

$$0\leqslant B(x-y,x-y)$$

$$= B(x,x) - B(x,y) - B(y,x) + B(y,y)$$

因此

$$B(x, y) + B(y, x) \leqslant B(x, x) + B(y, y).$$

从而对所有的 $\lambda \in \langle 0, 1\rangle$

$$B(\lambda x + (1-\lambda)y,$$
$$\lambda x + (1-\lambda)y) \leqslant \lambda B(x,x) + (1-\lambda)B(y,y)\ .$$

若对所有的 $x \in V$，$x \neq 0$，$B(x,x) > 0$，则函数 $x \to B(x, x)$ 是严格凸的。

满足上述条件之一的双线性函数例子是 V 的内积，和在 $\mathrm{R}^n \times \mathrm{R}^n$ 上的函数 $(x,y) \to (Ax \mid y)$，其中 A 是对称的 $n \times n$ 矩阵，是半正定或正定的。

(b) 设 E 是线性赋范空间．C 是 E 的非空凸子集，$x_0 \in E$．x_0 到 C 的距离定义为

$$\mathrm{d}(x_0, C) := \inf_{y \in C} \|x_0 - y\|$$

我们有

$$\mathrm{d}(x_0, C) = \inf_{y \in E} \{\|x_0 - y\| + \delta_C(y)\}$$
$$= (f \square \delta_C)(x_0)$$

其中 $f(x) = \|x\|$（见§5.15和§5.17）。由此推得函数 $x \to \mathrm{d}(x, C)$ 是凸的，建议读者对此命题构造一个更为直接的证明。

(c) 在 (b) 中，我们用到函数 $x \to \|x\|$ 的凸性。一般说来，若 $C \subset V$ 是凸代数体，$0 \in C^i$，则 C 的度规（见§2.19）是 V 上的凸函数。

(d) 设 f：$V \to \overline{\mathrm{R}}$ 是凸的且 $a \in V$。则

$$(f \square \delta_{\{a\}})(x) = f(x-a) \qquad (x \in V)$$

(e) V上最简单的凸函数是线性函数和仿射函数．V上的仿射函数形如

$$x \longmapsto f(x)+\alpha \tag{34}$$

这里f：$V\to \mathbf{R}$ 是线性的，$\alpha\in\mathbf{R}$．E上的连续仿射函数具有(34)式的形式，其中的f是E上的连续线性函数．

5.19 方向导数

设f：$V\to\overline{\mathbf{R}}$，$x_0$，$x\in V$．$f$在点$x_0$沿方向$x$的方向导数定义为极限

$$f'(x_0; x) := \lim_{\varepsilon\downarrow 0}\frac{f(x_0+\varepsilon x)-f(x_0)}{\varepsilon},$$

如果它存在（$+\infty$，$-\infty$允许作为极限）．

定理． 设f是V上的正常凸函数，$x_0\in\mathrm{dom}(f)$．则

(a) 对任何$x\in V$，$f'(x_0; x)$存在．

(b) V上的函数$x\longmapsto f'(x_0; x)$是正齐次的和凸的．

证明． (a) 设$x\in V$．定义g：$\mathbf{R}\to\overline{\mathbf{R}}$为 $g(\varepsilon)=f(x_0+\varepsilon x)$ $(\varepsilon\in\mathbf{R})$．$g$是正常凸函数，因此 $g'_+(0)$存在(见§1.24)．由$g'_+(0)=f'(x_0; x)$，命题得证．

(b) 由方向导数的定义直接得知，对所有的$x\in V$，$\lambda\geqslant 0$，有 $f'(x_0; \lambda x)=\lambda f'(x_0; x)$，由此函数$x\longmapsto f'(x_0; x)$是正齐次的．该函数的凸性从下面推知

$$\frac{1}{\varepsilon}[f(x_0+\varepsilon(\lambda y+(1-\lambda)z)-f(x_0)]$$

$$=\frac{1}{\varepsilon}[f(\lambda(x_0+\varepsilon y)+(1-\lambda)(x_0+\varepsilon z))-f(x_0)]$$

$$\leqslant \frac{1}{\varepsilon}[\lambda f(x_0+\varepsilon y)+(1-\lambda)f(x_0+\varepsilon z)-f(x_0)]$$

$$=\lambda\cdot\frac{f(x_0+\varepsilon y)-f(x_0)}{\varepsilon}+(1-\lambda)\cdot\frac{f(x_0+\varepsilon z)-f(x_0)}{\varepsilon}$$

其中y，$z\in V$，$\varepsilon>0$，$\lambda\in\langle 0,1\rangle$．

连 续 性

5.20

$f: E\to \mathrm{R}$称为在点$a\in E$是局部上（下）有界的，如果存在a的一个邻域，f在此邻域是上（下）有界的。

定理. 设$f: E\to\overline{\mathrm{R}}$是凸的，若$a\in E$，使$f(a)>-\infty$且$f$在点$a$是局部上有界的，则

(a) f是正常凸函数，且$\mathrm{int}(\mathrm{dom}(f))\neq\phi$。

(b) f在$\mathrm{int}(\mathrm{dom}(f))$的每点是局部上有界的。

(c) f在$\mathrm{int}(\mathrm{dom}(f))$上连续。

证明. 对f的研究等价于对函数$x\longmapsto f(x+a)$的研究。因此，不失一般性，我们假设$a=0$．设U是0的邻域，在此邻域上f上有界：对任何$x\in U$，$f(x)\leqslant M<+\infty$。

(a) 我们有$U\subset\mathrm{int}(\mathrm{dom}(f))$。因为$f(0)>-\infty$，由§5.12推得$f$是正常凸函数。

(b) 设$x_0\in\mathrm{int}(\mathrm{dom}(f))$，存在$\lambda>1$，使得$\lambda x_0\in\mathrm{int}(\mathrm{dom}(f))$。

集合

$$W := x_0 + \left(1 - \frac{1}{\lambda}\right) U$$

是 x_0 的邻域。如果 $y \in W$，对某个 $u \in U$，我们有

$$y = x_0 + \left(1 - \frac{1}{\lambda}\right) u。$$

因此

$$f(y) = f\left(\frac{1}{\lambda}(\lambda x_0) + \left(1 - \frac{1}{\lambda}\right) u\right)$$

$$\leqslant \frac{1}{\lambda} f(\lambda x_0) + \left(1 - \frac{1}{\lambda}\right) f(u)$$

$$\leqslant \frac{1}{\lambda} f(\lambda x_0) + \left(1 - \frac{1}{\lambda}\right) M.$$

从而 f 在 x_0 局部上有界。

(c) 设 $0 < \varepsilon < 1$。集 $X := \varepsilon(U \cap (-U))$ 是 0 的邻域，对所有的 $x \in X$，我们有 $x/\varepsilon \in U$，所以 $f(x/\varepsilon) \leqslant M$ 及

$$f(x) = f\left((1-\varepsilon)\cdot 0 + \varepsilon \cdot \frac{1}{\varepsilon} x\right)$$

$$\leqslant (1-\varepsilon) f(0) + \varepsilon f\left(\frac{1}{\varepsilon} x\right)$$

$$\leqslant f(0) + \varepsilon[M - f(0)].$$

我们还有 $-x/\varepsilon \in U$，故

$$f(0) = f\left(\frac{1}{1+\varepsilon} x + \frac{\varepsilon}{1+\varepsilon}\left(-\frac{1}{\varepsilon} x\right)\right)$$

$$\leqslant \frac{1}{1+\varepsilon} f(x) + \frac{\varepsilon}{1+\varepsilon} f\left(-\frac{1}{\varepsilon} x\right)$$

$$\leqslant \frac{1}{1+\varepsilon} f(x) + \frac{\varepsilon}{1+\varepsilon} M.$$

从而 $|f(x)-f(0)| \leqslant \varepsilon(M-f(0))$。即 f 在 0 连续，结合此结果与 (b)，(c) 即成立。

5.21

现设 E 为线性赋范空间。这时，可证得更多的结果。

设 $A \subset E$，f：$E \to \overline{\mathbf{R}}$ 称为相对于 A 是李普西兹的，如果 f 在 A 上为实值函数，且存在 $K>0$，使对任何的 x，$y \in A$，成立

$$|f(x)-f(y)| \leqslant K\|x-y\|.$$

f：$E \to \overline{\mathbf{R}}$ 称为相对于 $A \subset E$ 为局部李普西兹的，如果 f 是 A 上的实值函数，且对每个 $a \in A$，存在 a 的邻域 U，使 f 相对于 $U \cap A$ 是李普西兹的。

定理. 设 E 为线性赋范空间，f：$E \to \overline{\mathbf{R}}$ 是正常凸函数。若 f 在 E 的某点是局部上有界的，则 f 相对于 $\operatorname{int}(\operatorname{dom}(f))$ 是局部李普西兹的（见§1.7）。

证明. 设 $a \in \operatorname{int}(\operatorname{dom}(f))$。从§5.20 得知 f 在 a 连续。因此存在 $r_0>0$ 及 m，$M \in \mathbf{R}$，使对闭球 $B(a;\ r_0)=\{x \in E \mid \|x-a\| \leqslant r_0\}$ 中的所有点 x，有 $m \leqslant f(x) \leqslant M$。设 $0<r<r_0$ 及 x，$y \in B(a;\ r)$。记 $\|x-y\|=\sigma$，$z=y+[(r_0-r)/\sigma] \cdot (y-x)$，我们有 $y=\lambda z+(1-\lambda)x$。这里 $\lambda=\sigma/(\sigma+r_0-r)$。由于 $\|z-a\| \leqslant \|y-a\|+r_0-r \leqslant r_0$，必 $z \in B(a;\ r_0)$。从而 $f(y) \leqslant \lambda f(z)+(1-\lambda)f(x)$，因此

$$
\begin{aligned}
f(y)-f(x) &\leqslant \lambda(f(z)-f(x)) \\
&\leqslant \lambda(M-m) \\
&\leqslant \frac{M-m}{r_0-r}\|x-y\| .
\end{aligned}
$$

交换 x 和 y，得

$$
|f(x)-f(y)| \leqslant \frac{M-m}{r_0-r}\|x-y\|
$$

对所有 x，$y \in B(a, r)$。证明完成。

R 中的连续性及下半连续性

5.22 引理

设 $A \subset \mathbf{R}^n$ 是开集，$x_0 \in A$。则存在一个 n- 单纯形 S，使得 $S \subset A$ 且 $x_0 \in \operatorname{int}(S)$。

证明。存在 $\varepsilon>0$，使闭球 $B(x_0, \varepsilon)$ 含于 A 中。设 e_1, e_2, …, e_n 是 $\mathbf{R}^n$ 的标准正交基，令

$$
P:=-\frac{\varepsilon}{2n}\sum_{i=1}^{n} e_i
$$

n-单纯形 $T:=\operatorname{co}\left(\frac{1}{2}\varepsilon e_1, \frac{1}{2}\varepsilon e_2, \cdots, \frac{1}{2}\varepsilon e_n, P\right)$ 含于 $B\left(0;\right.$ $\left.\frac{1}{2}\varepsilon\right)$ 中。原点 0 可以表示成

$$
0=\sum_{i=1}^{n} \frac{1}{2n}\left(\frac{1}{2}\varepsilon e_i\right)+\frac{1}{2} P=\sum_{i=1}^{n} \mu_i\left(\frac{1}{2}\varepsilon e_i\right)+\mu_0 P
$$

其中 $0<\mu_i<1$，$(0\leqslant i\leqslant n)$ 且 $\sum_{i=0}^{n}\mu_i=1$。由此 $0\in\text{int}(T)$，$x_0\in\text{int}(T+x_0)$。由于 $T+x_0$ 是含于 A 中的 n-单纯形，引理得证。

5.23 定理

设 f：$\mathbb{R}^n\to\overline{\mathbb{R}}$ 为凸函数，则 f 在 $\text{ri}(\text{dom}(f))$ 上的限制是连续的。

证明。若 f 为非正常凸函数，则 f 在 $\text{ri}(\text{dom}(f))$ 上的限制是常值函数 $-\infty$（见§5.12）。因此这个限制是连续的。

若 f 为正常凸函数，存在 $x_0\in\text{ri}(\text{dom}(f))$。如果 $\dim(\text{ri}(\text{dom}(f)))=k$，由引理 5.22，存在一个 k- 单纯形 $S=\text{co}(a_0,a_1,\cdots,a_k)$，使 $S\subset\text{ri}(\text{dom}(f))$ 且 $x_0\in\text{ri}(S)$。设 $x\in S$，$x=\sum_{i=0}^{k}\lambda_i a_i$，其中 $\lambda_i\geqslant 0$ $(0\leqslant i\leqslant k)$，$\sum_{i=0}^{k}\lambda_i=1$。我们有

$$f(x)\leqslant\sum_{i=0}^{k}\lambda_i f(a_i)\leqslant\max_{0\leqslant i\leqslant k} f(a_i)\ .$$

因此，f 在 $\text{aff}(\text{dom}(f))$ 上的限制在 x_0 是局部上有界的。应用§5.20即得所述结果。

注。

(a) 从上述定理推出 $\mathbb{R}^n$ 上的实凸函数是连续的。

(b) 上述定理不能推出凸函数 $\mathbb{R}^n\to\overline{\mathbb{R}}$ 在 $\text{ri}(\text{dom}(f))$ 中每点连续（建议读者构造一个反例）。

但是，凸函数 $\mathbb{R}^n\to\overline{\mathbb{R}}$ 确实在 $\text{ri}(\text{dom}(f))$ 的每点下半连续（这个简单的证明留给读者）。

5.24 定理

设 f 是 $\mathbf{R}^n$ 上的正常凸函数，则 $\bar{f}$ 是正常凸函数。

证明．设 $x_0 \in \mathrm{ri}(\mathrm{dom}(f))$。我们有 $f(x_0) \in \mathbf{R}$。由 § 5.22 的注（b），f 在 x_0 下半连续，因此 $\bar{f}(x_0) = f(x_0)$（见定理 5.8）。用 § 5.12 断定函数 $\bar{f}$（$\bar{f}$ 是凸的和下半连续的）不可能是非正常的。

5.25

由 § 5.21和 § 5.23，得到下面结果：

定理． 设 f 为 $\mathbf{R}^n$ 上的正常凸函数，则 f 相对于 $\mathrm{ri}(\mathrm{dom}(f))$ 是局部李普西兹的。

5.26

设 E 为线性赋范空间，$A \subset E$。令 $T := \{f_\beta \mid \beta \in B\}$ 是函数 $E \to \overline{\mathbf{R}}$ 的集族。T 称为相对于 A 是局部等度李普西兹的，如果每个 $f_\beta \in T$ 是 A 上的实值函数，对每个 $a \in A$，存在 a 的邻域 U 及 $K > 0$，使得对所有的 x，$y \in U \cap A$ 和所有的 $\beta \in B$ 有

$$|f_\beta(x) - f_\beta(y)| \leqslant K \|x - y\|.$$

定理． 设 $U \subset \mathbf{R}^n$ 是开的，$T := \{f_\beta \mid \beta \in B\}$ 是凸函数 $\mathbf{R}^n \to \overline{\mathbf{R}}$ 的集族。若对每个 $x \in U$，集合 $\{f_\beta(x) \mid \beta \in B\}$ 是有界的（换言之，T 在 U 上点态有界），则 T 相对于 U 是局部等度李普西兹的。

证明（见定理5.23的证明）。设 $a\in U$，定义

$$M(x):=\sup\{f_\beta(x)\mid \beta\in B\}\qquad (x\in U)$$

$$m(x):=\inf\{f_\beta(x)\mid \beta\in B\}\qquad (x\in U).$$

由引理5.22，存在 n-单纯形 $S=\mathrm{co}(a_0, a_1, \cdots, a_n)$，使得 $S\subset U$ 且 $a\in \mathrm{int}(S)$。有

$$(\forall\beta\in B)(\forall x\in S)\ f_\beta(x)\leqslant M,$$

这里$M:=\max\{M(a_i)\mid 0\leqslant i\leqslant n\}$。

设 $r_0>0$，使得 $B(a; r_0)\subset S$，假定 $x\in B(a; r_0)$。令 $\|x-a\|=\sigma$，$y=a+(r_0/\sigma)(a-x)$，就有 $y\in B(a; r_0)$ 及 $a=\lambda x+(1-\lambda)y$，这里 $\lambda=r_0/(r_0+\sigma)$。从而对所有的 $\beta\in B$

$$f_\beta(a)\leqslant\lambda f_\beta(x)+(1-\lambda)f_\beta(y)\leqslant\lambda f_\beta(x)+(1-\lambda)M.$$

因此

$$f_\beta(x)\geqslant\frac{1}{\lambda}f_\beta(a)+\frac{\lambda-1}{\lambda}M\geqslant m$$

其中 $m:=\min(M, m(a))$。我们得

$$(\forall\beta\in B)(\forall x\in B(a; r_0))\quad m\leqslant f_\beta(x)\leqslant M.$$

设 $0<r<r_0$，遵循§5.21中的证明，得出

$$|f_\beta(x)-f_\beta(y)|\leqslant\frac{M-m}{r_0-r}\|x-y\|$$

对任何 x，$y\in B(a; r)$，$\beta\in B$。所述结果得证。

注。U 的开性假定可以减弱，此证留给读者。

5.27

请读者注意，线性函数的某些性质与凸函数类似。例如，定理5.23对 $\mathbb{R}^n$ 上每个实线性函数亦真。§5.26可对应成定理：Bamach 空间上的实线性连续函数集族若点态有界，则必

是一致有界（即存在$K>0$，对集族中每个函数f，有$\|f\|\leqslant K$）。

可 微 凸 函 数

5.28 定义

设E为线性赋范空间，E'是E的对偶（见§3.13）。设f：$E\to\overline{\mathbf{R}}$，$x_0\in E$，$f$在$x_0$有限。

(a) f称为在x_0是Frechet-可微(或简称可微)，如果存在$x'\in E'$，使得

$$\lim_{x\to x_0}\frac{f(x)-f(x_0)-\langle x-x_0\mid x'\rangle}{\|x-x_0\|}=0 \qquad (35)$$

或

$$f(x)=f(x_0)+\langle x-x_0\mid x'\rangle+o(\|x-x_0\|)\quad (x\to x_0)$$

x'由（35）式唯一确定。它被称为f在x_0的Frechet导数(或简称为导数)，记成$f'(x_0)$或$df(x_0)$。

(b) f称为在x_0是Gateaux—可微，如果存在$x'\in E'$，使得对所有的$x\in E$

$$\lim_{\varepsilon\to 0}\frac{f(x_0+\varepsilon x)-f(x_0)}{\varepsilon}=\langle x\mid x'\rangle \qquad (36)$$

x'由（36）式唯一确定。称x'为f在x_0的Gateaux-微分，用$\nabla f(x_0)$表示。

Frechet—可微性蕴涵Gateaux—可微性，但其逆不真。然对$\mathbf{R}^n$上的正常凸函数，却是真的。

如果Gateaux—微分$\nabla f(x_0)$存在，则对每个$x\in E$有$f'(x_0, x)$存在，此处

$$f'(x_0; x)=\langle x|\nabla f(x_0)\rangle.$$

但f在x_0的所有方向导数的存在性不蕴涵f在x_0的Gateaux—可微性（因为后一性质意味着$f'(x_0; x)$除对每个$x\in E$存在外，而且关于x是线性和连续的）.

5.29 定理

设f：$\mathbf{R}^n\to\mathbf{R}$具有连续的一阶和二阶偏导数（换言之，$f$是二次连续可微的）. f在x处的Hesse矩阵$H(x)$记为

$$H(x):=\begin{bmatrix}\frac{\partial^2 f}{\partial x_1^2}(x) & \cdots & \frac{\partial^2 f}{\partial x_1\partial x_n}(x)\\ \vdots & & \vdots\\ \frac{\partial^2 f}{\partial x_n\partial x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x)\end{bmatrix}.$$

其中$x\in\mathbf{R}^n$. 则

$$f\text{是凸的}\Longleftrightarrow\text{对}\ \forall x\in\mathbf{R}^n,\ H(x)\text{是半正定的}.$$

证明. 由§5.9知f是凸的$\Longleftrightarrow$对所有x，$y\in\mathbf{R}^n$，从R到R的函数g：$t\to f(x+ty)$是凸的. 由定理1.11，g是凸的当且仅当对所有的$t\in\mathbf{R}$，$g''(t)\geqslant 0$. 现有

$$g'(t)=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(x+ y)y_i$$

$$g''(t)=\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_j}(x+ty)y_iy_j=(H(x+ty)y|y).$$

因此

$$f\text{是凸的}\Longleftrightarrow\text{对所有的}x,\ y\in\mathbf{R}^n,\ t\in\mathbf{R}$$
$$(H(x+ty)y\,|\,y)\geqslant 0$$

所述结果得证。

次 可 微 性

设E是线性赋范空间，E'为E的对偶。

5.30

凸函数并非是必定可微的。下面引入次可微性。将会看到，在凸分析中凸函数的次梯度往往是有用的，如果寻常的导数并不存在。

定义。设$f: E\to\overline{\mathbf{R}}$，$x_0\in E$，$f$在$x_0$有限。

(a) 设$x_0'\in E'$。x_0'称为是f在x_0的次梯度，如果对任何的$x\in E$，

$$f(x)\geqslant f(x_0)+\langle x-x_0\mid x_0'\rangle .$$

(b) f在x_0的所有次梯度的集合叫做f在x_0的次微分，记为$\partial f(x_0)$。$\partial f(x_0)$是E'的凸子集。f称为在x_0是次可微的，如果$\partial f(x_0)\neq\phi$。若f在x不是有限，定义$\partial f(x)=\phi$。

(c) f的次微分是从E到E'的多值函数$\partial f: x\longmapsto\partial f(x)$。

(d) ∂f的有效域$\mathrm{dom}(\partial f)$是集$\{x\in E\mid\partial f(x)\neq\phi\}$。

5.31 例

(a) 设$f: \langle a, b\rangle\to\mathbf{R}$是凸的。由定理1.6，$f$在每点$c\in\langle a, b\rangle$是次可微的，且有

$$\partial f(c)=[f_-'(c),\ f_+'(c)].$$

(b) 定义$f: \mathbf{R}\to\mathbf{R}$为

$$f(x)=\begin{cases}-\sqrt{1-x^2} & |x|\leqslant 1\\ +\infty & |x|>1\end{cases}$$

f 是 $\langle -1, 1\rangle$ 上次可微（甚至可微）的正常凸函数。-1，$1\in\mathrm{dom}(f)$，但 f 在 -1，或 $+1$ 不是次可微的。

(c) 设 $E=\mathbf{R}^n$，f 是 $\mathbf{R}^n$ 上的欧氏范数：$f(x)=\|x\|$。f 在原点 0 不可微，但却是次可微。$\partial f(0)$ 由所有这样的点 $x'\in\mathbf{R}^n$ 组成

$$(\forall y\in\mathbf{R}^n)\ \|y\|\geqslant(y\mid x')。$$

显见，$\partial f(0)$ 是闭单位球 $\{x\in\mathbf{R}^n \mid \|x\|\leqslant 1\}$。

5.32

下面给出次可微性的几何描述。首先忆及每个 $F\in(E\oplus\mathbf{R})'$ 可被表成

$$F(x, \lambda)=\langle x\mid x_0'\rangle+\alpha_0\lambda \qquad ((x, \lambda)\in E\oplus\mathbf{R})$$

对某个 $x_0'\in E'$ 和某个 $\alpha_0\in\mathbf{R}$。记此为 $F=(x_0', \alpha_0)$。

因而，在 $E\oplus\mathbf{R}$ 中的闭超平面 H 是由形如

$$\langle x\mid x_0'\rangle+\alpha_0\lambda=\beta_0 \tag{37}$$

的方程描述的集，这里 $(x_0', \alpha_0)\neq(0, 0)$ 且 $\beta_0\in\mathbf{R}$。H 称为是垂直的，如果 $\alpha_0=0$（因此，$x_0'\neq 0$）。若 H 非垂直（这意味着 $\alpha_0\neq 0$），则（37）式可写成

$$\lambda=\left\langle x\,\middle|\,-\frac{1}{\alpha_0}x_0'\right\rangle+\frac{\beta_0}{\alpha_0}$$

由此得知 H 是从 E 到 $\mathbf{R}$ 的连续仿射函数

$$x\longmapsto\left\langle x\,\middle|\,-\frac{1}{\alpha_0}x_0'\right\rangle+\frac{\beta_0}{\alpha_0}$$

的图象。

5.33 引理

设 f 为 E 上的正常凸函数，$x_0 \in \mathrm{dom}(f)$。假定 $F := (x_0', \alpha_0) \in (E \oplus \mathrm{R})'$，$H := F^{-1}(\beta_0)$ 为 $\mathrm{epi}(f)$ 在 $(x_0, f(x_0))$ 的支撑超平面。则

(a) 如果 $F(\mathrm{epi}(f)) \geqslant \beta_0$，那么 $\alpha_0 \geqslant 0$。

(b) 如果 $x_0 \in \mathrm{int}(\mathrm{dom}(f))$，那么 $\alpha_0 \neq 0$（换言之，H 是非垂直的）。

证明. (a) 由 $F(\mathrm{epi}(f)) \geqslant \beta_0$。对所有的 $(x, \lambda) \in \mathrm{epi}(f)$，有

$$\langle x \mid x_0' \rangle + \alpha_0 \lambda \geqslant \beta_0 .$$

令 $\lambda \to +\infty$，得出 $\alpha_0 \geqslant 0$。

(b) 设 $x_0 \in \mathrm{int}\,(\mathrm{dom}(f))$。若有 $\alpha_0 = 0$，$F(\mathrm{epi}(f)) \geqslant \beta_0$，则

$$\langle x \mid x_0' \rangle \geqslant \beta_0 = \langle x_0 \mid x_0' \rangle .$$

于是，对所有的 $x \in \mathrm{dom}(f)$ 有

$$\langle x - x_0 \mid x_0' \rangle \geqslant 0 . \tag{38}$$

因为 $x_0 \in \mathrm{int}(\mathrm{dom}(f))$，对每个 $y \in E$，存在 $\varepsilon > 0$，使 $x_0 + \varepsilon y \in \mathrm{dom}(f)$，$x_0 - \varepsilon y \in \mathrm{dom}(f)$。在 (38) 式中，分别令 $x = x_0 + \varepsilon y$ 及 $x = x_0 - \varepsilon y$，可得 $\langle y \mid x_0' \rangle \geqslant 0$ 和 $\langle y \mid x_0' \rangle \leqslant 0$，因此 $\langle y \mid x_0' \rangle = 0$。既然对所有的 $y \in E$，有 $\langle y \mid x_0' \rangle = 0$，则必 $x_0' = 0$，矛盾。

5.34 定理

设 f 为 E 上的正常凸函数，$x_0 \in \mathrm{dom}(f)$。则 f 在 x_0 是次可微的当且仅当 $\mathrm{epi}(f)$ 在 $(x_0, f(x_0))$ 处存在非垂直的闭支

撑超平面。

证明。必要性：设$x_0' \in \partial f(x_0)$。令$F := (x_0', -1)$，$\beta_0 := \langle x_0 \mid x_0' \rangle - f(x_0)$，$H := F^{-1}(\beta_0)$。$H$是$E \oplus \mathbf{R}$中非垂直闭超平面。我们有$F(x_0, f(x_0)) = \beta_0$，从而$(x_0, f(x_0)) \in H$。设$(x, \lambda) \in \mathrm{epi}(f)$。我们有$f(x) \leqslant \lambda$及$f(x) \geqslant f(x_0) + \langle x - x_0 \mid x_0' \rangle$，故

$$
\begin{aligned}
F(x, \lambda) &= \langle x \mid x_0' \rangle - \lambda \\
&\leqslant \langle x \mid x_0' \rangle - f(x) \\
&\leqslant \langle x_0 \mid x_0' \rangle - f(x_0) = \beta_0
\end{aligned}
$$

即H为$\mathrm{epi}(f)$在$(x_0, f(x_0))$的支撑超平面。

充分性：设$H = F^{-1}(\beta_0)$是$\mathrm{epi}(f)$在$(x_0, f(x_0))$的非垂直闭支撑超平面。对某个$x_0' \in E'$及某个$\alpha_0 \in \mathbf{R}$，我们有$F = (x_0', \alpha_0)$。不失一般性，假定

$$F(\mathrm{epi}(f)) \geqslant \beta_0 \tag{39}$$

其中$\beta_0 = F(x_0, f(x_0))$。结合(39)式及引理5.33得到$\alpha_0 \geqslant 0$。因$\alpha_0 \neq 0$，必$\alpha_0 > 0$。由(39)式，对每个$x \in \mathrm{dom}(f)$

$$\langle x \mid x_0' \rangle + \alpha_0 f(x) \geqslant \beta_0 = \langle x_0 \mid x_0' \rangle + \alpha_0 f(x_0)$$

因此

$$f(x) \geqslant f(x_0) + \left\langle x - x_0 \,\middle|\, -\frac{1}{\alpha_0} x_0' \right\rangle. \tag{40}$$

由于$x \notin \mathrm{dom}(f)$时，(40)式平凡地成立，因此$-x_0'/\alpha_0 \in \partial f(x_0)$，$\partial f(x_0) \neq \phi$。

5.35 定理

设f为E上的正常凸函数。

(a) 如 f 在某点 $x_0\in\mathrm{dom}(f)$ 连续，则 f 在 $\mathrm{int}(\mathrm{dom}(f))$ 中每点是次可微的。

(b) 如 $E=\mathbf{R}^n$，则 f 在 $\mathrm{ri}(\mathrm{dom}(f))$ 的每点是次可微的。

证明. (a) 因 f 在 x_0 连续，故存在 x_0 的邻域 U，在 U 上 f 上有界：存在 $k>0$，使 $f(U)\leqslant k$。从而 $U\times\langle k,+\infty\rangle\subset\mathrm{epi}(f)$，因此 $\mathrm{epi}(f)$ 是 $E\oplus\mathbf{R}$ 中的凸体。由§3.10，$\mathrm{epi}(f)$ 在 $(x_0,f(x_0))$ 存在一个非平凡的闭支撑超平面 H，因 $x_0\in\mathrm{int}(\mathrm{dom}(f))$，从引理5.33知 H 是非垂直的。再用定理5.34，得出 $\partial f(x_0)\neq\phi$。

由§5.20，f 在 $\mathrm{int}(\mathrm{dom}(f))$ 上连续。按上面结果，对任意 $x\in\mathrm{int}(\mathrm{dom}(f))$，必有 $\partial f(x)\neq\phi$。

(b) 设 $x_0\in\mathrm{ri}(\mathrm{dom}(f))$。我们把证明 $(x_0,f(x_0))\in\mathrm{rb}(\mathrm{epi}(f))$ 留给读者。由§5.12，$\mathrm{epi}(f)$ 在 $(x_0,f(x_0))$ 存在非平凡的支撑超平面 H。设 H 是垂直的。$H=F^{-1}(\beta_0)$，其中 $F=(x_0',0)$。遵循引理5.33的证明，对所有的 $y\in\mathrm{dom}(f)$，有

$$\langle y\mid x_0'\rangle=\beta_0$$

因此，$\mathrm{epi}(f)\subset H$，这与 H 是 $\mathrm{epi}(f)$ 的非平凡支撑超平面事实矛盾。故 H 必是非垂直的，因而 $\partial f(x_0)\neq\phi$。

推论. 凸函数 $f:\mathbf{R}^n\to\mathbf{R}$ 是次可微的。

5.36 定理

设 $f:E\to\overline{\mathbf{R}}$ 为凸函数，$x_0\in E$，且 f 在 x_0 有限，则 $x_0'\in\partial f(x_0)$ 当且仅当对所有的 $x\in E$

$$f'(x_0;x)\geqslant\langle x\mid x_0'\rangle\ .$$

证明。 有

$$x_0' \in \partial f(x_0)$$
$$\Longleftrightarrow (\forall x\in E)\, f(x) \geqslant f(x_0) + \langle x - x_0 \mid x_0'\rangle$$
$$\Longleftrightarrow (\forall x\in E)\,(\forall \varepsilon>0)\, f(x_0+\varepsilon x) \geqslant f(x_0) + \varepsilon\langle x \mid x_0'\rangle$$
$$\Longleftrightarrow (\forall x\in E)\,(\forall \varepsilon>0)\, \frac{1}{\varepsilon}\{f(x_0+\varepsilon x) - f(x_0)\} \geqslant \langle x \mid x_0'\rangle$$

因为

$$\frac{1}{\varepsilon}\{f(x_0+\varepsilon x) - f(x_0)\} \downarrow f'(x_0;\ x) \qquad (\varepsilon \downarrow 0)$$

（见§1.5），所述结果得证。

5.37 定理

设 f：$E \to \overline{\mathrm{R}}$ 为凸函数，f 在 $x_0 \in E$ 有限，若 f 在 x_0 为 G—可微，则

$$\partial f(x_0) = \{\nabla f(x_0)\}\ .$$

其中 $\nabla f(x_0)$ 是 f 在 x_0 的 G—微分.

证明. 从§5.28推出，对所有的 $x \in E$，

$$f'(x_0;\ x) = \langle x \mid \nabla f(x_0)\rangle\ .$$

应用定理5.36得 $\nabla f(x_0) \in \partial f(x_0)$. 反之，如果 $x_0' \in \partial f(x_0)$，则由定理5.36，有

$$\langle x \mid \nabla f(x_0)\rangle \geqslant \langle x \mid x_0'\rangle$$

因此，对所有的 $x \in E$，

$$\langle x \mid \nabla f(x_0) - x_0'\rangle \geqslant 0\ ,$$

于是 $\nabla f(x_0) - x_0' = 0$，故 $x_0' = \nabla f(x_0)$.

注. 如 f 在 x_0 连续，且在 x_0 有唯一的次梯度，可证 f 在 x_0 是 G— 可微.

5.38 定理

设 f_1，f_2 为E上的正常凸函数，则

(a) 对所有的 $x \in E$，

$$\partial(f_1+f_2)(x) \supset \partial f_1(x) + \partial f_2(x) .$$

(b) 如存在 $\mathrm{dom}(f_1) \cap \mathrm{dom}(f_2)$ 中一个点，使 f_1 在此点连续，则对所有的 $x \in E$，

$$\partial(f_1+f_2)(x) = \partial f_1(x) + \partial f_2(x) .$$

证明. (a) 若 $x'_1 \in \partial f_1(x)$，$x'_2 \in \partial f_2(x)$，则对所有的 $y \in E$有

$$\begin{cases} f_1(y) \geqslant f_1(x) + \langle y-x \mid x'_1 \rangle \\ f_2(y) \geqslant f_2(x) + \langle y-x \mid x'_2 \rangle \end{cases}$$

因此

$$(f_1+f_2)(y) \geqslant (f_1+f_2)(x) + \langle y-x \mid x'_1 + x'_2 \rangle .$$

从而 $x'_1 + x'_2 \in \partial(f_1+f_2)(x)$，(a) 得证。

(b) 设 $x_0 \in \mathrm{dom}(f_1+f_2)$ （$=\mathrm{dom}(f_1) \cap \mathrm{dom}(f_2)$），$x'_0 \in \partial(f_1+f_2)(x_0)$（我们把$x_0 \in E$，$x_0 \bar{\in} \mathrm{dom}(f_1+f_2)$ 的情形留给读者证明）. 定义 g_1 和 g_2 为

$$\begin{cases} g_1(x) := f_1(x+x_0) - f_1(x_0) - \langle x \mid x'_0 \rangle & (x \in E) \\ g_2(x) := f_2(x+x_0) - f_2(x_0) & (x \in E) \end{cases} \tag{41}$$

g_1 和 g_2 是E上的正常凸函数，且 $\mathrm{dom}(g_1) = \mathrm{dom}(f_1) - x_0$，$\mathrm{dom}(g_2) = \mathrm{dom}(f_2) - x_0$，$g_1(0) = g_2(0) = 0$，$0 \in \partial(g_1+g_2)(0)$。而且，$g_1$ 在 $\mathrm{dom}(g_1) \cap \mathrm{dom}(g_2)$ 中某点连续。令

$$C_1 := \{(x, \lambda) \in E \oplus \mathbf{R} \mid g_1(x) \leqslant \lambda\}$$
$$C_2 := \{(x, \lambda) \in E \oplus \mathbf{R} \mid \lambda \leqslant -g_2(x)\}$$

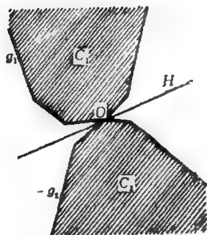


图 12

C_1 和 C_2 是非空凸集（注意 $C_1=\mathrm{epi}(g_1)$）（见图12）.

鉴于 g_1 在 $\mathrm{dom}(g_1)$ 中某点的连续性，可有 $\mathrm{int}(C_1)\neq\phi$（见定理 5.35 的证明），且 $\mathrm{int}(C_1)\cap C_2=\phi$。事实上，若此结果不真，则存在 $(x,\ \lambda)\in\mathrm{int}(C_1)\cap C_2$ 及 $\varepsilon>1$，使 $\{x\}\times\langle\lambda-\varepsilon,\ \lambda+\varepsilon\rangle\subset C_1$，因此 $g_1(x)\leqslant\lambda-\varepsilon<\lambda\leqslant-g_2(x)$，这将导致矛盾，因 $0\in\partial(g_1+g_2)(0)$，蕴涵对所有的 $x\in E$ 成立 $g_1(x)\geqslant-g_2(x)$。

由定理 3.8，存在 $E\oplus\mathrm{R}$ 中的闭超平面 H 真分离 C_1 和 C_2。设 $H=F^{-1}(\beta)$，$F(C_1)\leqslant\beta$，$F(C_2)\geqslant\beta$，$F=(x',\ \alpha)$，这里 $(x',\ \alpha)\neq(0,\ 0)$。因 $(0,\ 0)\in H$，我们有 $\beta=0$。假设 α 为 0（因此 $x'\neq0$），则有

$$\langle x\mid x'\rangle\leqslant0\qquad\forall x\in\mathrm{dom}(g_1)\tag{42}$$

$$\langle x\mid x'\rangle\geqslant0\qquad\forall x\in\mathrm{dom}(g_2)\tag{43}$$

设 $y\in\mathrm{dom}(g_1)\cap\mathrm{dom}(g_2)$，$g_1$ 在 y 连续，则 $y\in\mathrm{int}(\mathrm{dom}(g_1))$。(42) 式蕴涵 $\langle y\mid x'\rangle<0$，这与 (43) 式相矛盾。故 $\alpha\neq0$。

对每个 $\lambda>0$，$(0,\ \lambda)\in C_1$，我们有 $\alpha<0$。令 $y_0'=-x'$

$/\alpha$，就有

$$\begin{cases} \langle x \mid y_0' \rangle \leqslant g_1(x) & \text{对所有的 } x \in \mathrm{dom}(g_1) \\ \langle x \mid y_0' \rangle \geqslant -g_2(x) & \text{对所有的 } x \in \mathrm{dom}(g_2) \end{cases}$$

利用（41）式，对所有的 $x \in E$ 有

$$\begin{cases} f_1(x+x_0) \geqslant f_1(x_0) + \langle x \mid x_0' + y_0' \rangle \\ f_2(x+x_0) \geqslant f_2(x_0) + \langle x \mid -y_0' \rangle \end{cases}$$

因此，$x_0' + y_0' \in \partial f_1(x_0)$，$-y_0' \in \partial f_2(x_0)$。我们得出

$$x_0' = (x_0' + y_0') + (-y_0') \in \partial f_1(x_0) + \partial f_2(x_0).$$

（b）证毕。

注。 $\partial(f_1+f_2)(x) = \partial f_1(x) + \partial f_2(x)$ 一般不真。

5.39

在有限维情形，可以证得一个较强的结果。

定理。 设 f_1、f_2 为 $\mathbf{R}^n$ 上的正常凸函数，满足 $\mathrm{ri}(\mathrm{dom}(f_1)) \cap \mathrm{ri}(\mathrm{dom}(f_2)) \neq \phi$，则对任意的 $x \in \mathbf{R}^n$，成立

$$\partial(f_1+f_2)(x) = \partial f_1(x) + \partial f_2(x).$$

证明。 溯及定理 5.38(b) 的证明。由 g_1，g_2，C_1，C_2 的定义，就有 $\mathrm{ri}(C_1) \cap \mathrm{ri}(C_2) = \phi$。根据 § 4.11，存在 $\mathbf{R}^n$ 中超平面 $H = F^{-1}(\beta)$ 真分离 C_1 和 C_2，设 $F = (x', \alpha)$，若 $\alpha = 0$，则集合

$$\{x \in \mathbf{R}^n \mid \langle x \mid x' \rangle = 0\}$$

为 $\mathbf{R}^n$ 中真分离 $\mathrm{dom}(g_1)$ 和 $\mathrm{dom}(g_2)$ 的超平面，这将引起矛盾。因为 $\mathrm{ri}(\mathrm{dom}(f_1)) \cap \mathrm{ri}(\mathrm{dom}(f_2)) \neq \phi$ 蕴涵 $\mathrm{ri}(\mathrm{dom}(g_1)) \cap$

ri(dom(g_2))$\neq\phi$（见§4.11）。故必$\alpha\neq 0$．

定理后部分证明类于定理5.38(b)．

练　　习

1　见§5.9．证明

$f|m$是凸的$\Longleftrightarrow f_m$是凸的．

2　设V和W是（R上的）线性空间．T：$V\to W$是线性映射，f：$W\to\overline{\mathrm{R}}$是凸函数．定义$g$：$V\to\overline{\mathrm{R}}$为$g(x)=f(Tx)$．证明$g$是凸的．

3　设E为线性拓扑空间，凸函数f：$E\to\overline{\mathrm{R}}$，其有效域是相对开的．证明

$(\forall x\in E)f(x)>-\infty$或者$(\forall x\in E)\ |f(x)|=+\infty$．

4　证明§5.15中的定理．

5　设V和W是线性空间．f：$V\times W\to\overline{\mathrm{R}}$是凸的．定义$g$：$V\to\overline{\mathrm{R}}$为

$$g(x)=\inf\{f(x,\ y)|\ y\in W\}.$$

证明g是凸的．

6　设V是线性空间，$A\subset V\times\mathrm{R}$是凸集．定义$f$：$V\to\overline{\mathrm{R}}$为

$$f(x)=\inf\{\lambda\mid(x,\ \lambda)\in A\}.$$

证明f是凸的．

7　设E为线性拓扑空间，f是E上的正常凸连续函数，$\lambda\in\mathrm{R}$．令$A:=\{x\in E\mid f(x)\leqslant\lambda\}$，$B:=\{x\in E\mid f(x)<\lambda\}$．证明下述命题：

(a)　$B=\mathrm{int}(A)$未必真．

(b) 如存在 $x\in E$，使 $f(x)<\lambda$，则 $B=\operatorname{int}(A)$。

8 设 V 是线性空间，f 为 V 上的实凸函数。证明下述命题：

(a) f 的每个局部极小点必是整体极小点。

(b) 若 f 严格凸，则 f 至多有一个整体极小点。

(见第一章练习1)。

9 设 E 为内积空间。$C\subset E$ 是非空凸集，$x_0\in E$。证明至多存在一点 $c\in C$，使得 c 最接近于 x_0，即

$$\|x_0-c\|=\mathrm{d}(x_0,\ C)$$

(见§5.18，例(b))。

10 设 V 是线性空间，f，g：$V\to\overline{\mathbf{R}}$ 均为凸函数，请确定 $\operatorname{dom}(f\square g)$。

11 设 E 为线性拓扑空间。f：$E\to\overline{\mathbf{R}}$ 为凸函数，$a\in\operatorname{dom}(f)$。证明 f 在 a 连续当且仅当 f 在 a 局部上有界。

12 函数 f：$\mathbf{R}^n\to\overline{\mathbf{R}}$ 定义为

$$f(\xi_1,\xi_2,\cdots,\xi_n)=\begin{cases}-(\xi_1\xi_2\cdots\xi_n)^{1/n} & \text{若 } \xi_1\geqslant 0,\ \xi_2\geqslant 0,\ \cdots,\xi_n\geqslant 0\\ +\infty & \text{否则。}\end{cases}$$

证明 f 是凸的。

13 设 f：$\mathbf{R}^n\times\mathbf{R}\to\mathbf{R}$ 为连续函数，使得对每个 $t\in[a,b]$，函数 $x\longmapsto f(x,t)$ 是凸的(简言之：f 关于 x 是凸的)。定义函数 g 为

$$g(x)=\int_a^b f(x,t)\,\mathrm{d}t$$

证明 g 是凸的。

14 设 E 为线性赋范空间。证明从 E 到 $\mathbf{R}$ 的次可微函数是

凸的。（见第一章练习 3）。

15 设函数 f 定义为 R^n 上的 **Tchebycheff** 范数，即

$$f(\xi_1,\xi_2,\cdots,\xi_n)=\max_i|\xi_i|.$$

证明 $\partial f(0)=\mathrm{co}(e_1,\cdots,e_n,-e_1,\cdots,-e_n)$，这里 $e_1,e_2,\cdots,e_n$ 为 R^n 中的单位向量（$e_1=(1,0,0,\cdots,0)$，等等）。

16 设 E 为线性赋范空间，假定 $f(x)=\frac{1}{2}\|x\|^2\ (x\in E)$。试证

(a) f 是凸的。

(b) 对所有的 $x\in E$ 成立

$$\partial f(x)=\{x'\in E'\mid\langle x\mid x'\rangle=\|x\|\cdot\|x'\|\text{且}\|x'\|=\|x\|\}$$

这里

$$\|x'\|:=\sup_{x\neq 0}\frac{|\langle x\mid x'\rangle|}{\|x\|}$$

(c) 若 E 为 Hilbert 空间，则对 $\forall x\in E$，$\partial f(x)=\{x\}$。

17 设 E 为线性赋范空间，$f: E\to R$ 为实 G—可微，具有 G—微分 ∇f。证明下述命题：

(a) f 是凸的，当且仅当对所有的 x，$y\in E$，有

$$f(y)\geqslant f(x)+\langle y-x\mid\nabla f(x)\rangle.$$

(b) f 是严格凸的，当且仅当对所有的 x，$y\in E$，$x\neq y$ 有

$$f(y)>f(x)+\langle y-x\mid\nabla f(x)\rangle.$$

18 借助于一个反例说明

$$\partial(f_1+f_2)(x)=\partial f_1(x)+\partial f_2(x)$$

一般不成立（见定理5.38后面的注）。

19　证明在定理5.38中通过令$E=\mathbf{R}^n$得到的定理是§5.39中定理的一种特殊情形。

虽然下述练习是基本的，有用的，却有点儿缺乏趣味。

20　设X是拓扑空间，$A\subset X$。$f: A\to\overline{\mathbf{R}}$。证明下述条件等价：

(a)　对每个$\lambda\in\mathbf{R}$，集合$\{x\in A\mid f(x)\leqslant\lambda\}$是$X$的闭子集。

(b)　$\{(x,\lambda)\in A\times\mathbf{R}\mid f(x)\leqslant\lambda\}$是$X\times\mathbf{R}$的闭子集。

(c)　f是下半连续的，且对每个$a\in\overline{A}\setminus A$

$$\lim_{x\to a} f(x)=+\infty.$$

21　(a)　设E是线性拓扑空间，$C\subset E$为凸集，$f: C\to\mathbf{R}$为实凸函数。证明或者

(Ⅰ)　f在$\overline{C}$的每点是局部下有界的，或者

(Ⅱ)　对$\overline{C}$中所有的点，f均不局部下有界。

注。 f称为在$a\in\overline{C}$是局部下有界的，如果存在a的邻域U，使得f在$U\cap C$上是下有界的。

(b)　设$C\subset\mathbf{R}^n$是凸的，$f: C\to\mathbf{R}$为实凸函数。证明f在$\overline{C}$的每点是局部下有界的。

(c)　求E，C和f，使(a)的(Ⅱ)成立。

22　设$U\subset\mathbf{R}^n$为非空开凸集，$f: U\to\mathbf{R}$是凸的。证明f在U的每个紧子集上是李普西兹的（见§1.7）。

23　设f是凸函数$\mathbf{R}^p\times\mathbf{R}^q\to\mathbf{R}$，而且$f(x, y)$关于$x$连

续，关于y是凸的（即对每个$y\in R^p$，$x\longmapsto f(x, y)$是连续函数，对每个$x\in R^p$，$y\longmapsto f(x, y)$是凸函数）。证明f是连续的。

24 设E为线性赋范空间，凸函数$f: E\to \overline{R}$在点x_0连续，且$f(x_0)$有限。证明：对所有的$x\in E$，有

$$f'(x_0; x)=\max\{\langle x|x_0'\rangle \mid x_0'\in\partial f(x_0)\}.$$

25 证明定理5.37后面的'注'。

注 释

1 可以证明定义在 Banach 空间上的凸下半连续函数在 int(dom(f)) 中连续。（见§5.20）。

2 W.W. Breckner and G. Orban, Continuity properties of rationally s-convex mappings with values in an ordered topological linear space, 'Babes-Bolyai' University of Cluj-Napoca, 1978

在上文中凸函数的连续性（见§5.20）结果已被推广到取值于序线性拓扑空间的映射f，这样的f对任何x，y和任意$\lambda\in\langle 0, 1\rangle$，适合不等式

$$f(\lambda x+(1-\lambda)y)\leqslant\lambda^s f(x)+(1-\lambda)^s f(y),$$

其中s是属于区间$\langle 0, 1]$的实数。

3 R^n中的凸分析是我们课题的最经典和最有进展部分。一本基础参考书是

R.T. Rockafeller, *Convex Analysis*, Princeton, Princeton University Press, 1970。

4 可微性的更详尽的叙述见

A. D. Ioffe and V. M. Tihomirov, *Theory of Extremal Problems*, Amsterdam, North-Holland, 1979.

5 可以证明若 f 是 Banach 空间上的下半连续正常凸函数，则 $\mathrm{dom}(\partial f)\neq\phi$（见§6.35）且 $\mathrm{dom}(\partial f)$ 甚至在 $\mathrm{dom}(f)$ 中是稠密的。例如，这可见

I. Ekeland and R. Teman, *Convex Analysis and Variational Problems*, Amsterdam, North-Holland, 1976.

6 定理5.38在下列文献中可找到

R. T. Rockafellar, *Convex functions and dual extremum problems*, PhD Thesis, Harvard University, 1963.

第六章 对偶性

本章我们用 E 表示 R 上的线性赋范空间（不止含一点），具有范数 $x\to\|x\|$，用 E' 表示E 的对偶。分离定理推出对每个 $x\in E$，$x\neq 0$，存在 $x'\in E'$，使得$\langle x\,|\,x'\rangle\neq 0$。

共 轭 函 数

6.1 定义

(a) $f: E\to\overline{\mathrm{R}}$ 的共轭（或者对偶，或若极）$f^*: E'\to\overline{\mathrm{R}}$，定义为

$$f^*(x')=\sup_{x\in E}\{\langle x\,|\,x'\rangle-f(x)\}\qquad (x'\in E').$$

(b) $g: E'\to\overline{\mathrm{R}}$ 的共轭 $g^*: E\to\overline{\mathrm{R}}$ 定义为

$$g^*(x)=\sup_{x'\in E'}\{\langle x\,|\,x'\rangle-g(x')\}\qquad (x\in E).$$

(c) 从 E 到 $\overline{\mathrm{R}}$ 或从 E' 到 $\overline{\mathrm{R}}$ 的函数 f 的二次极函数 f^{**} 是 f 的共轭的共轭 $(f^*)^*$。

6.2 注。

(a) 若 $f^*(x')$有限，则它是满足不等式

$$f(x) \geqslant \langle x | x' \rangle - \alpha \qquad \forall x \in E$$

的最小实数 α（见1.16）。

(b) 每个形如 $x \to \langle x | x' \rangle + \alpha$ 的函数是 E 上的连续仿射函数，这里 $x' \in E'$，$\alpha \in R$（因此，$x \to \langle x, x' \rangle - g(x')$亦是 E 上的连续仿射函数）。而 E 上的每个连续仿射函数是这种形式。

如果 E' 具有用范数 $x' \to \| x' \|$ 定义的范拓扑，这里

$$\| x' \|_* = \sup_{x \neq 0} \frac{|\langle x | x' \rangle|}{\| x \|} = \sup_{\| x \| = 1} |\langle x | x' \rangle|$$

那么，每个形如 $x' \to \langle x | x' \rangle + \alpha$ 的函数是 E' 上的连续仿射函数，其中 $x \in E$，$\alpha \in R$（因此，$x' \to \langle x | x' \rangle - f(x)$亦是 E' 上的连续仿射函数）。但E'上的连续仿射函数一般不具备这种形式。例如，如果我们赋 E' 以所谓的弱拓扑$W(E', E)$（在这种拓扑下的收敛性是指对序列—(x'_α)，或者更一般地，网—(x'_α)收敛到点 x' 意味着对每个 $x \in E$，$\langle x | x'_\alpha \rangle \to \langle x | x' \rangle$）便是这种情形。

6.3

下述简单性质的证明留给读者。

(a) 若 $f: E \to \overline{R}$，$h: E \to \overline{R}$，且 $f \leqslant h$，则 $f^* \geqslant h^*$。

(b) $(+\infty)^* = -\infty$

(c) 设 $f: E \to \overline{R}$，若 f 在某点取值为$-\infty$，则 $f^* = +\infty$。特别地，$(-\infty)^* = +\infty$。

注意，由(b)和(c)可推得 $f^{**} = f$，但此式一般不成立。读者易证对所有的 $f: E \to \overline{R}$，我们有 $f^{**} \leqslant f$。

(d) 设 $\{f_\alpha | \alpha \in A\}$ 为任意多个 $E \to \overline{R}$ 函数的集族，则

$$(\inf_\alpha f_\alpha)^* = \sup_\alpha f_\alpha^*$$

$$(\sup_\alpha f_\alpha)^* \leqslant \inf_\alpha f_\alpha^*.$$

后一个不等式中，等号通常不成立。

(e) 设f：$E\to\overline{\mathbf{R}}$，$\lambda>0$，则

$$(\lambda f)^*(x')=\lambda f^*\left(\frac{x'}{\lambda}\right)\qquad(x'\in E').$$

(f) 设f：$E\to\overline{\mathbf{R}}$，$\alpha\in\mathbf{R}$，则

$$(f+\alpha)^*=f^*-\alpha.$$

(g) 设f：$E\to\overline{\mathbf{R}}$，$x\in E$，$x'\in E'$，则

$$f_x^*(x')=f^*(x')+\langle x|x'\rangle,$$

这里函数f_x定义为$f_x(y)=f(y-x)\qquad(y\in E)$。

(h) 设f：$E\to\overline{\mathbf{R}}$，则

$$\inf\{f(x)\,|\,x\in E\}=-f^*(0).$$

6.4 例

(a) 设$x_0'\in E'$，$\alpha\in\mathbf{R}$。定义f：$E\to\mathbf{R}$为

$$f(x)=\langle x|x_0'\rangle-\alpha\qquad(x\in E).$$

则

$$f^*(x')=\sup_{x\in E}\left\{\langle x|x'-x_0'\rangle+\alpha\right\}=\begin{cases}+\infty & x'\neq x_0'\\ \alpha & x'=x_0'\end{cases}$$

因此$f^*=\delta_{\{x_0'\}}+\alpha$(见§5.15)。

(b) 设$f(x)=\frac{1}{2}\|x\|^2$。则

$$f^*(x')=\sup_{x\in E}\left\{\langle x|x'\rangle-\frac{1}{2}\|x\|^2\right\}$$

$$=\sup_{t\geq 0}\sup_{\|x\|=t}\left\{\langle x|x'\rangle-\frac{1}{2}\|x\|^2\right\}$$

$$= \sup_{t \ge 0} \left\{ t \sup_{\|x\|=1} \langle x | x' \rangle - \frac{1}{2} t^2 \right\}$$

$$= \sup_{t \ge 0} \left\{ t \|x'\| - \frac{1}{2} t^2 \right\} = \frac{1}{2} \|x'\|^2$$

因此 $f^*(x) = \frac{1}{2} \|x'\|^2$。

6.5 定义

设$A \subset F$。A 的支撑函数是 A 的指标函数 δ_A 的共轭 δ_A^*：

$$\delta_A^*(x') = \sup_{x \in E} \{ \langle x | x' \rangle - \delta_A(x) \}$$
$$= \sup_{x \in A} \langle x | x' \rangle \qquad (x' \in E').$$

我们建议读者就 $E = \mathbf{R}^n$ 的情形，给出 δ_A^* 的几何解释。

6.6 定理

设f：$E \to \mathbf{R} \cup \{+\infty\}$，$g$：$E \to \mathbf{R} \cup \{+\infty\}$。

则

$$(f \Box g)^* = f^* + g^*$$

证明. 对每 $x' \in E'$，我们有

$$(f \Box g)^*(x') = \sup_x \{ \langle x | x' \rangle - \inf_{x_1 + x_2 = x} [f(x_1) + g(x_2)] \}$$

$$= \sup_x \sup_{x_1 + x_2 = x} [\langle x | x' \rangle - f(x_1) - g(x_2)]$$

$$= \sup_{x_1, x_2} \{ [\langle x_1 | x' \rangle - f(x_1)] + [\langle x_2 | x' \rangle - g(x_2)] \} = f^*(x') + g^*(x').$$

6.7 例

设 $C\subset E$ 为非空凸集。定义 $f: E\to\mathbf{R}$ 为

$$f(x)=\inf_{y\in C}\|x-y\| \qquad (x\in E).$$

从 § 5.18的例(b)得出 $f=g\square\delta_C$，这里 $g(x)=\|x\|(x\in E)$。对每个 $x'\in E'$，我们有

$$\begin{aligned} g^*(x') &= \sup_x\{\langle x|x'\rangle-\|x\|\} \\ &= \sup_{t\geq 0}\sup_{\|x\|=t}\{\langle x|x'\rangle-t\} \\ &= \sup_{t\geq 0} t(\|x'\|-1)=\begin{cases} +\infty & 若\|x'\|>1 \\ 0 & 若\|x'\|\leqslant 1\end{cases}\end{aligned}$$

因此，$g^*=\delta_S$，这里 $S=\{x'\in E'\mid \|x'\|\leqslant 1\}$。从而

$$f^*=g^*+\delta_C^*=\delta_S+\delta_C^*$$

所以

$$f^*(x')=\begin{cases}\delta_C^*(x') & 若\|x'\|\leqslant 1 \\ +\infty & 若\|x'\|>1.\end{cases}$$

6.8 定理

设 $f: E\to\overline{\mathbf{R}}$ 则 f^* 是 E' 上的下半连续凸函数（它具有范拓扑；见 § 6.2的（b））。

这个简单的证明留给读者。

6.9

设 $f: E\to\overline{\mathbf{R}}$，对每个 $x\in E$，$x'\in E'$，

$$f^*(x')\geqslant\langle x|x'\rangle-f(x)$$

从而

$$f(x)+f^*(x')\geqslant\langle x|x'\rangle \tag{44}$$

其中左边无论甚么时候都是有定义的。(44) 式称为 Fenchel 不等式。(见 § 1.16)

6.10 定理

设 $f: E\to\overline{\mathbf{R}}$, 且 f 在 $x\in E$ 有限。则

$$x'\in\partial f(x)\Longleftrightarrow f^*(x')=\langle x|x'\rangle-f(x). \tag{45}$$

证明. 我们有

$$\begin{aligned} x'\in\partial f(x) &\Longleftrightarrow(\forall y\in E)f(y)\geqslant f(x)+\langle y-x|x'\rangle\\ &\Longleftrightarrow(\forall y\in E)\langle x|x'\rangle-f(x)\geqslant\langle y|x'\rangle-f(y)\\ &\Longleftrightarrow\sup_{y\in E}\{\langle y|x'\rangle-f(y)\}=\langle x|x'\rangle-f(x)\\ &\Longleftrightarrow f^*(x')=\langle x|x'\rangle-f(x)\end{aligned}$$

注. 最后的等式可以写成

$$f(x)+f^*(x')=\langle x|x'\rangle.$$

由此可见 f 的次梯度是 E' 中这样的元素，它使 Fenchel 不等式成为等式。

二 次 极 函 数

6.11 定理

设 $f: E\to\overline{\mathbf{R}}$. 则

(a) f^{**} 是 f 的所有连续仿射弱函数集族的点态上确界。

(b) f^{**} 是 E 上的下半连续凸函数。

(c) $f^{***}=f^*$.

证明. (a) 设 A 为 f 的所有连续仿射弱函数的集族. 令

$F=\sup\{g\mid g\in A\}$。

首先假设对某个 $x'\in E$，$f^*(x')=-\infty$。读者易证这时有 $f^{**}=f=F=+\infty$，从而 $f^{**}=F$。其次，假设对每个$x'\in E'$，$f^*(x')>-\infty$。则对每个满足 $f^*(x')<+\infty$的 $x'\in E'$，函数 g：$x\to\langle x|x'\rangle-f^*(x')$ 是 E 上的连续仿射函数。从§6.9得知 g 是 f 的弱函数，故 $g\in A$。从而对所有的 $x\in E$

$$f^{**}(x)=\sup\{\langle x|x'\rangle-f^*(x')\mid f^*(x')<+\infty\}\leqslant F(x).$$

若 $h\in A$，形如 $h(x)=\langle x|x'\rangle-\alpha$ $(x\in E)$这里 $x'\in E'$，$\alpha\in\mathbf{R}$，我们有

$$(\forall x\in E)\langle x|x'\rangle-\alpha\leqslant f(x)$$

因此

$$\alpha\geqslant\sup_{x\in E}\{\langle x|x'\rangle-f(x)\}=f^*(x')$$

于是

$$h(x)=\langle x|x'\rangle-\alpha\leqslant\langle x|x'\rangle-f^*(x').$$

从而，对每个 $x\in E$

$$F(x)\leqslant\sup_{x'\in E'}\{\langle x|x'\rangle-f^*(x')\}=f^{**}(x).$$

由之得出 $f^{**}=F$。

(b) 见定理 6.8。

(c) 我们有 $f^{**}\leqslant f$（见§6.3(c)），从而 $f^{***}\geqslant f^*$（见§6.3(a)）。在不等式 $f^{**}\leqslant f$ 中，用 f^* 代替 f，即得 $f^{***}\leqslant f^*$，这样证明了所述结果。

6.12 引理

设 g：$E\to\overline{\mathbf{R}}$ 是下半连续的正常凸函数，则 g 有连续仿射弱函数。更确切地说，对每个 $x_0\in E$ 和每个 $\alpha_0\in\langle-\infty,$

$g(x_0)\rangle$，存在一个连续仿射函数 h，使得 $h(x_0)=\alpha_0$，$h<g$。

证明. 设 $x_0\in E$，$-\infty<\alpha<g(x_0)$．我们有 (x_0, α_0) $\overline{\in}$ $\mathrm{epi}(g)$．因为 $\mathrm{epi}(g)$ 是 $E\oplus\mathrm{R}$ 的非空闭凸子集，由定理 3.12 知存在 $E\oplus\mathrm{R}$ 中一个闭超平面 $H=F^{-1}(\beta)$ 严格分离 $\mathrm{epi}(g)$ 和 (x_0, α_0)．假设 $F=(x', \alpha)$ 且

$$(\forall(x, \lambda)\in\mathrm{epi}(g))\langle x|x'\rangle+\alpha\lambda>\beta \tag{46}$$

$$\langle x_0|x'\rangle+\alpha\alpha_0<\beta \tag{47}$$

在 (46) 中令 $\lambda\to+\infty$，即得 $\alpha\geqslant0$。

(a) 若 $g(x_0)<+\infty$，(46) 与 (47) 式推出

$$\langle x_0|x'\rangle+\alpha g(x_0)>\beta>\langle x_0|x'\rangle+\alpha\alpha_0$$

因此 $\alpha(g(x_0)-\alpha_0)>0$，从而 $\alpha>0$。定义 E 上的连续仿射函数 h 为

$$h(x)=\left\langle x-x_0\middle|-\frac{1}{\alpha}x'\right\rangle+\alpha_0 \qquad (x\in E).$$

由 (46) 和 (47) 式，对 $\forall x\in\mathrm{dom}(g)$ 有

$$g(x)>\frac{\beta}{\alpha}+\left\langle x\middle|-\frac{1}{\alpha}x'\right\rangle$$

$$=h(x)+\left\langle x_0\middle|-\frac{1}{\alpha}x'\right\rangle+\frac{\beta}{\alpha}-\alpha_0>h(x)$$

因此 $h<g$，$h(x_0)=\alpha_0$。

(b) 若 $g(x_0)=+\infty$和$\alpha>0$，我们可以给出一个与(a) 的后半部证明类似的证明。

(c) 现在假定 $g(x_0)=+\infty$，$\alpha=0$。定义 E 上的连续仿射函数 h 为

$$h(x)=\langle x|-x'\rangle+\beta \qquad (x\in E).$$

我们有 $h(x_0)>0$ 及对 $\forall x\in \mathrm{dom}(g)$，$k(x)<0$。从(a)得知 g 有连续仿射弱函数 m，使得 $m<g$。若 $m(x_0)\geqslant\alpha_0$，则函数 $h:=m+\alpha_0-m(x_0)$ 满足 $h(x_0)=\alpha_0$，$h<g$。若 $m(x_0)<\alpha_0$，对每个 $\lambda>0$，函数 $m+\lambda k$ 是满足 $m+\lambda k<g$ 的连续仿射函数。定义 $h:=m+\lambda_0 k$，这里

$$\lambda_0=\frac{\alpha_0-m(x_0)}{k(x_0)}$$

我们又有 $h(x_0)=\alpha_0$，$h<g$。

推论。一个下半连续的正常凸函数是它的连续仿射弱函数集族的点态上确界。若 g 是这样的函数，则 $g^{**}=g$（见定理6.11(a)）。

6.13

为了给出二次极函数的其他（简单）特性，我们引入函数的闭包概念，它与下半连续包的概念密切相关。

定义。设 $f: E\to\overline{\mathbf{R}}$。

(a) f 的闭包 $\mathrm{cl}(f)$ 定义为 f 的下半连续包 $\bar f$，如果 $\bar f$ 处处不等于 $-\infty$。而在另外的情形，它定义为常值 $-\infty$。

(b) f 称为闭的，如果 $\mathrm{cl}(f)=f$。

6.14 例

(a) 设 $f: E\to\overline{\mathbf{R}}$ 是凸的。如果 $\bar f$ 在某点取值 $-\infty$，由§5.12 $\bar f$ 无有限值，且

$$\bar f(x)=\begin{cases}-\infty & \text{如果 } x\in\overline{\mathrm{dom}(f)}\\ +\infty & \text{如果 } x\notin\overline{\mathrm{dom}(f)}\end{cases}$$

在这种情况下，$\mathrm{cl}(f)$ 仅在 $\overline{\mathrm{dom}(f)}$ 之外与 $\bar f$ 不同，因 $\mathrm{cl}(f)$

为 $-\infty$，而 $\bar{f}$ 为 $+\infty$。

(b) 若 f 是 R^n 上的正常凸函数，则由定理5.24，$\bar{f}$ 也是正常凸的。这时有 $\mathrm{cl}(f)=\bar{f}$。因此，对 R^n 上的正常凸函数，闭性与下半连续性相同。

6.15 定理

设 $f: E\to\overline{\mathrm{R}}$。则

$$f^{**}=\mathrm{cl}(\mathrm{co}(f))。$$

证明。 令 $g:=\mathrm{cl}(\mathrm{co}(f))$。$g$ 是下半连续的凸函数。

(a) 若 $g=+\infty$，则 $f=+\infty$，因此 $f^{**}=+\infty=g$。

(b) 假定 $g=-\infty$。若 f 有连续仿射弱函数 h。则 $h\leqslant f$。因此 $h\leqslant\overline{\mathrm{co}(f)}$。这蕴涵对每个 $x\in E$，$\overline{\mathrm{co}(f)}(x)>-\infty$，因此 $\overline{\mathrm{co}(f)}=\mathrm{cl}(\mathrm{co}(f))=g$，矛盾。由此导出 f 没有连续仿射弱函数，所以 $f^{**}=-\infty=g$。

(c) 剩下的情况，g 是下半连续正常凸函数，且 $g=\overline{\mathrm{co}(f)}$。我们有 $f^{**}\leqslant g$。若有 $x_0\in E$，使得 $f^{**}(x_0)<g(x_0)$。则存在 $\alpha\in\mathrm{R}$，使 $f^{**}(x_0)<\alpha<g(x_0)$。由引理6.12，存在连续仿射函数 h，使得 $h(x_0)=\alpha$ 且 $h<g$。这推出 $f^{**}(x_0)\geqslant h(x_0)=\alpha$，矛盾。就得 $f^{**}=g$。

6.16 定理

若 f 为 R^n 上的正常凸函数，则 $f^{**}=\bar{f}$。

证明。 结合定理6.15和§6.14的例 (b) 即得。

注。见定理1.15.

集合 $\Gamma(E)$

6.17 定义

$\Gamma(E)$ 是所有这种函数$E\to\overline{\mathbf{R}}$ 的集族，它们是E上形为$x\to\langle x|x'\rangle+\alpha$ 函数族的点态上确界，这里 $x'\in E'$，$\alpha\in\mathbf{R}$。

类似地，定义 $\Gamma(E')$ 是所有这种函数 $E'\to\overline{\mathbf{R}}$ 的集族，它们是 E' 上形如 $x'\to\langle x|x'\rangle+\alpha$ 的函数族的点态上确界，其中 $x\in E$，$\alpha\in\mathbf{R}$。注意，$\Gamma(E)$亦能定义成所有这种函数 $E\to\overline{\mathbf{R}}$ 的集族，它们是 E 上连续仿射函数族的点态上确界。$\Gamma(E')$ 也可类似地定义，只要我们赋 E' 以适当的拓扑（例如，弱拓扑；见§6.2，注(b)）。

6.18 定理

设 f：$E\to\overline{\mathbf{R}}$。下列条件是等价的：

(a) $f\in\Gamma(E)$。

(b) $f=f^{**}$。

(c) f 是下半连续函数正常凸，或者 f 是常值函数 $-\infty$ 和 $+\infty$ 之一。

证明. (a) $\Rightarrow$(c)：设 $f\in\Gamma(E)$。因 f 是连续仿射（因此凸）函数族 A 的点态上确界，故 f 是下半连续的凸函数。若 $A=\Phi$，则 $f=-\infty$。若$A\neq\Phi$，则对每个 $x\in E$，$f(x)>-\infty$。因此 $f=+\infty$，或者 f 是正常凸函数。

(c) $\Rightarrow$(b)：由§6.3 性质 (b) 和 (c)，有 $(-\infty)^{**}=$

$-\infty$及 $(+\infty)^{**}=+\infty$。若f是下半连续的正常凸函数，由定理6.15有 $f^{**}=\mathrm{cl}(f)=\bar{f}=f$ （见引理6.12的推论）。

(b)⇒(a)：利用定理6.11(a)。

6.19 定理

映射 $f\to f^*$ 是 $\Gamma(E)$ 和 $\Gamma(E')$ 间的一个双射。

证明. 设 $f\in\Gamma(E)$。由 f^* 和 $\Gamma(E')$ 的定义推出 $f^*\in\Gamma(E')$。又 $f^{**}=f$ （见定理6.18），故 $\Gamma(E)$ 到 $\Gamma(E')$ 的映射 $f\to f^*$ 是单射。

设 $g\in\Gamma(E')$，则 $g^*\in\Gamma(E)$。遵循定理6.11(a)的证明，我们可证 g^{**} 是形如 $x'\to\langle x|x'\rangle+\alpha$ 的函数 g 之所有弱函数构成的族之点态上确界，这里 $x\in E$，$\alpha\in\mathbb{R}$。从而 $g\leqslant g^{**}\leqslant g$，因此，$g=g^{**}=(g^*)^*$。故从 $\Gamma(E)$ 到 $\Gamma(E')$ 的映射 $f\to f^*$ 是满射。这就证明了所述结果。

我们用 $\Gamma_0(E)$ 表示所有 $\Gamma(E)$ 中非常值函数 $+\infty$ 与 $-\infty$ 所组成的集。类似地可定义 $\Gamma_0(E')$。注意到 $\Gamma_0(E)$ 是 E 上所有下半连续正常凸函数的集族。由上面定理，映射 $f\to f^*$ 就是 $\Gamma_0(E)$ 与 $\Gamma_0(E')$ 之间的一个双射。

支 撑 函 数

6.20 定理

设 $C\subset E$ 是凸集，则：

(a) $C=\Phi\Longleftrightarrow\delta_C=+\infty$。

(b) $C\neq\Phi\Longleftrightarrow\delta_C$ 是正常凸的。

(c) C 是闭且非空的$\Longleftrightarrow \delta_C \in \Gamma_0(E)$.

(d) $\overline{\delta_C} = \delta_{\overline{C}}$。

(e) $\delta_C^{**} = \delta_{\overline{C}}$。

证明。(a)，(b) 和 (c) 是 § 5.15的直接结果。

(d) 若 $C = \Phi$ 等式$\overline{\delta_C} = \delta_{\overline{C}}$ 显然成立。若 $C \neq \Phi$，我们有

$$\text{epi}\,(\overline{\delta_C}) = \overline{\text{epi}(\delta_C)} = \overline{C \times [0, +\infty\rangle}$$

$$= \overline{C} \times [0, +\infty\rangle = \text{epi}\,(\delta_{\overline{C}}) \ ,$$

因此$\overline{\delta_C} = \delta_{\overline{C}}$。

(e) 若 $C = \Phi$，等式显然成立。若 $C \neq \Phi$，结合 (d) 与定6.15，我们有

$$\delta_C^{**} = \text{cl}(\delta_C) = \overline{\delta_C} = \delta_{\overline{C}}.$$

6.21 定理

设 $C \subset E$ 是闭的，凸的且非空。则 $\delta_C^* \in \Gamma_0(E')$.

证明。结合定理6.20 (c) 和 § 6.19。

6.22

从定义 6.5 得出，支撑函数是正齐次的。反之，我们有下面定理。

定理。设 $g \in \Gamma_0(E')$ 是正齐次的。则存在 E 的唯一非空闭凸子集 C，使得 $\delta_C^* = g$.

证明。首先我们证明 g^* 是一个指标函数。对每个 $\lambda > 0$和每个 $x' \in E'$，我们有

$$(\lambda g^*)^*(x')=\lambda g^{**}\left(\frac{x'}{\lambda}\right)=\lambda g\left(\frac{x'}{\lambda}\right)=g(x')$$

因此

$$(\lambda g^*)^*=g$$

于是

$$\lambda g^*=(\lambda g^*)^{**}=g^*.$$

故 g^* 的值仅是 0 和 $+\infty$。从而 $g^*=\delta_C$，这里

$$C=\{x\in E\mid g^*(x)=0\} \tag{48}$$

因此 $g=g^{**}=\delta_C^*$。读者易证 C 是非空闭凸集。现设 C_1 和 C_2 是 E 的非空闭凸子集，使得

$$\delta_{C_1}^*=\delta_{C_2}^*$$

则

$$\delta_{C_1}=\delta_{C_1}^{**}=\delta_{C_2}^{**}=\delta_{C_2}$$

从而

$$C_1=C_2.$$

注.

(a) 用 (48) 定义的集合 C 也可以写成

$$C=\{x\in E\mid(\forall x'\in E')\langle x|x'\rangle\leqslant g(x')\}.$$

(b) 上面定理推出，在 $\mathbb{R}^n$ 中，非空闭凸集与正齐次下半连续正常凸函数之间存在一一对应。

(c) 若 $C\subset E$，则 C 与 $\overline{C}$ 有相同的支撑函数，并它也是 E 中每个满足 $C\subset A\subset\overline{C}$ 的子集 A 的支撑函数。因此已知一个集的支撑函数而去确定这集，通常不可能。但上述定理指出，若我们知道这个集是凸和闭的，则将有可能。

练　　习

E 是线性赋范空间。

1 设 φ: $\mathrm{R}\to\overline{\mathrm{R}}$ 是偶函数。定义 f：$E\to\overline{\mathrm{R}}$ 和 g：$E'\to\overline{\mathrm{R}}$ 为 $f(x)=\varphi(\|x\|)$，$g(x')=\varphi^*(\|x'\|)$ $(x\in E,\ x'\in E')$ 证明 $f^*=g$。

2 证明满足等式 $f^*=f$ 的唯一函数 f：$\mathrm{R}^n\to\overline{\mathrm{R}}$ 是 $f(x)=\frac{1}{2}(x|x)$。（见§6.4，例(b)。）。

3 证明下述命题：若 f：$E\to\overline{\mathrm{R}}$ 有连续仿射弱函数，则 $f^{**}=\overline{\mathrm{co}(f)}$。

4 设 f：$E\to\overline{\mathrm{R}}$，$f$ 在 $x_0\in E$ 有限。$\partial f(x_0)\neq\Phi$。证明

(a) $f^{**}(x_0)=f(x_0)$；

(b) $\partial f^{**}(x_0)=\partial f(x_0)$。

5. 设 f：$E\to\overline{\mathrm{R}}$。证明 f^{**} 是 $\Gamma(E)$ 中满足 $g\leqslant f$ 的最大函数。

6. 设 f 是 R^n 上的下半连续正常凸函数。证明对所有的 x，$x'\in\mathrm{R}^n$

$$x'\in\partial f(x)\Longleftrightarrow x\in\partial f^*(x')\ .$$

7. 设 f：$\mathrm{R}^n\to\overline{\mathrm{R}}$ 为凸函数。证明

f 是正常凸的 $\Longleftrightarrow f^*$ 是正常凸的。

8. 设 $A\subset E$。证明 $\delta_A^{**}=\delta_B$，这里 $B=\overline{\mathrm{co}}(A)$。

9. 设 C、$D\subset E$ 是凸的。证明下述命题：

(a) $\delta_{C+D}^*=\delta_C^*+\delta_D^*$ （见定理6.6）

(b) $\overline{C}\subset\overline{D}\Longleftrightarrow\delta_C^*\leqslant\delta_D^*$。

10 证明 R^n 中有界非空凸集的支撑函数是实正齐次凸函数。

11 设 f: $\mathbf{R}^n \to \overline{\mathbf{R}}$ 是正齐次凸函数，且 $f \neq +\infty$。证明 $\mathrm{cl}(f)$ 是一个支撑函数。

12 设 f: $E \to \overline{\mathbf{R}}$ 为凸函数，f 在 $x_0 \in E$ 是有限的。设 g: $E \to \overline{\mathbf{R}}$ 定义为 $g(x) = f'(x_0; x) \quad (x \in E)$ 证明

$\partial f(x_0) = \partial g(0)$.

13 设 $K \subset E$ 是非空锥，证明

$\delta_K^* = \delta_{K^0}$。

14 设 $A \subset \mathbf{R}^n$ 使 δ_A^* 是线性函数。证明 A 由一个点组成。

注 释

1 最初，对凸函数的共轭作一般论述的文献是

W. Fenchel, On conjugate convex functions, *Can. J. Math.* 1 (1949) 73—7.

见第一章注 3。

2 凸函数的共轭概念与微分方程理论中用 $X = y'$, $Y = xy' - y$。所定义的经典勒让德变换有关。

3 共轭函数的定义是基于连续仿射弱函数集族的使用。（见§1.16和定义6.1）。利用其它弱函数类，我们可得到共轭性的各种推广。对这些推广提供途径的公理化对偶性方法已在下面的文献中给出。

J. J. M. Evers and H. van Maaren, Duality principles in mathematics and their relations to conjugate functions, Department of Applied Mathematics, Twente University of Technology, 1981.

第七章　最优化

多年来，特别在经典变分学中，最优化问题与可微性有关。虽然对凸函数已进行了长时间研究，第一章注1所提到的Jensen的论文是最初讨论实凸函数的文章之一，然而仅在现代，人们才发现凸函数在最优化中的广泛应用。原来对凸可微函数可以给出新的最优准则。进而，在去掉可微性，而考虑凸的、不可微的函数时，其中某些准则仍然是正确的。本章给出凸性在最优化中的作用的一个简述。

我们用E表示R上的线性赋范空间（不止含一点），用$x \longrightarrow \|x\|$表示范数，E'表示对偶。

7.1　设f是E上的正常凸函数，f在$x_0 \in E$有（整体）极小，当且仅当对每个$x \in E$

$$f(x) \geqslant f(x_0) = f(x_0) + \langle x - x_0 \mid 0 \rangle。$$

由此得到

$$f\text{在}x_0\text{有（整体）极小} \Longleftrightarrow 0 \in \partial f(x_0)。\qquad (49)$$

条件$0 \in \partial f(x_0)$有如下几何解释，epi(f)在$(x_0, f(x_0))$存在一个水平（闭）支撑超平面（见定理5.34）。这可看作是在凸函数中与可微函数取得极小的熟知条件的类似物，即在整体极小的情形下（见第五章练习8）f的图形的水平切平面是

$\mathrm{epi}(f)$ 的水平支撑超平面。

7.2

设$C\subset E$是非空凸集。f是E上的正常凸函数，使得$\mathrm{dom}(f)\cap C\neq\phi$。我们用 f_C 表示f在C上的限制。

令$g=f+\delta_C$；g是正常凸函数。在C上极小化f（即极小化f_C）等价于在E上极小化g，因此（49）式推出f_C在$x_0\in C$有极小，当且仅当$0\in\partial g(x_0)$。按照定理 5.38，在下述情形之一时，后一条件能写成

$$0\in\partial f(x_0)+\partial\delta_C(x_0). \qquad (50)$$

(a) 存在一个点属于$C\cap\mathrm{dom}(f)$，f在此点连续；

(b) $\mathrm{int}(C)\cap\mathrm{dom}(f)\neq\phi$.

7.3

下面将研究出现在（50）式中的集$\partial\delta_C(x_0)$。我们有

$$x'\in\partial\delta_C(x_0)\Longleftrightarrow(\forall x\in E)\delta_C(x)\geqslant\delta_C(x_0)+\langle x-x_0|x'\rangle$$
$$\Longleftrightarrow(\forall x\in C)\langle x-x_0|x'\rangle\leqslant 0 \qquad (51)$$

由此得知$\partial\delta_C(x_0)$是包含 0 的凸锥。我们把下述定理的简单证明留给读者。

定理。若$x_0\in\mathrm{int}(C)$，则$\partial\delta_C(x_0)=\{0\}$。

我们建议读者在$E=\mathbf{R}^n$的情形下，给出（51）式中最后那个不等式的几何解释（见图13）。$x'\in\partial\delta_C(x_0)$称为是$C$在$x_0$的法向量，$\partial\delta_C(x_0)$称为$C$在$x_0$的法锥（或支撑函数的锥）。

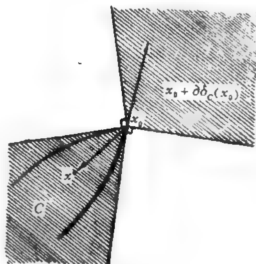


图13

7.4

在凸分析的应用中，我们会遇下述集合：

$$C=\{x\in R^n\,|\,g(x)\leqslant 0\} \tag{52}$$

其中 g 是R^n上的实凸函数，回忆起 g 将是连续和次可微的、且（见第五章练习 7）

$$\text{int}(C)=\{x\in R^n\,|\,g(x)<0\}\neq\phi$$

只要 g 满足所谓

Slater条件：存在 $x\in R^n$，使$g(x)<0$。

由§7.3，若$g(x_0)<0$，则$\partial\delta_C(x_0)=\{0\}$。下面将在$g(x_0)=0$的情形下研究集合$\partial\delta_C(x_0)$。

7.5 定理

设g是R^n上满足Slater条件的实凸函数，C是由（52）式定义的集合，假定$x_0\in R^n$，使$g(x_0)=0$。则

$$\partial\delta_C(x_0)=K^0$$

这里

$$K=\{x\in R^n | g'(x_0;\ x)<0\}.$$

证明。因为$\mathrm{int}(C)\neq\phi$，由定理2.27，有

$$C=\overline{\mathrm{int}(C)}$$

因此

$$x'\in\partial\delta_C(x_0)\Longleftrightarrow(\forall x\in\mathrm{int}(C))\langle x-x_0|\ x'\rangle\leqslant 0$$

（见(51)式）．对每个$x\in\mathrm{int}(C)$及每个$\lambda>0$成立

$$g'(x_0;\ \lambda(x-x_0))=\lambda\lim_{\varepsilon\downarrow 0}\frac{g(x_0+\varepsilon(x-x_0))-g(x_0)}{\varepsilon}$$

$$=\lambda\lim_{\varepsilon\downarrow 0}\frac{1}{\varepsilon}g((1-\varepsilon)x_0+\varepsilon x)$$

$$\leqslant\lambda g(x)<0.$$

反之，若$x\in R^n$，满足$g'(x_0;\ x)<0$，则存在$\varepsilon>0$，使$g(x_0+\varepsilon x)<0$，因此$x_0+\varepsilon x\in\mathrm{int}(C)$，故$x\in\Omega$（$\mathrm{int}(C)-x_0$），这里$\Omega=\langle 0,\ +\infty\rangle$。于是得知

$$K_{:}=\{x\in \mathbf{R}^n \mid g'(x_0;\ x)<0\}=\Omega(\mathrm{int}(C)-x_0).$$

所以

$$x'\in \partial\delta_C(x_0)\Longleftrightarrow(\forall x\in \mathrm{int}(C))\ \langle x-x_0\mid x'\rangle\leqslant 0$$

$$\Longleftrightarrow(\forall x\in K)\langle x\mid x'\rangle\leqslant 0$$

$$\Longleftrightarrow x'\in K^0.$$

就得$\partial\delta_C(x_0)=K^0$。

注：读者应注意在上面证明中，$\{x\mid g'(x_0;\ x)\leqslant 0\}=\Omega(C-x_0)$并非必然成立。

7.6 定理

设g：$\mathbf{R}^n\longrightarrow\mathbf{R}$是凸函数，$x_0\in\mathbf{R}^n$。定义函数$h$为

$$h(x)=g'(x_0;\ x)\qquad (x\in\mathbf{R}^n)$$

则h是$\partial g(x_0)$的支撑函数。

证明。由§5.19，h是正齐次和凸的。又因g是实凸函数，h只取有限值（见定理1.6），故h是连续函数，由§6.22，h是集

$$D_{:}=\{x\in\mathbf{R}^n\mid(\forall x'\in\mathbf{R}^n)(x\mid x')\leqslant g'(x_0;\ x)\}\tag{53}$$

的支撑函数。

由定理5.36，我们有

$$x\in\partial g(x_0)\Longleftrightarrow(\forall x\in R^n)g'(x_0;\ x')\geqslant(x'\mid x)\tag{54}$$

结合（53）和（54）式，得出所述结果。

注。请看第五章练习24。

7.7 引理

设 $A\supset R^n$ 是非空紧子集，$0\overline{\in}A$。令K是由A 生成的锥，即

$$K_{\ } = \{\lambda a \mid \lambda\geqslant 0,\ a\in A\}。$$

则K是闭的.

证明：设$x\in \overline{K}$，则存在非负实数列 (λ_n) 和 A 中的序列 (a_n)，使得 $\lambda_n a_n\to x$ $(n\to\infty)$. 因为 A 是紧的，于是存在 (a_n) 的子列 (b_n) 使 $b_n\to b\in A$. 设 (μ_n) 是 (λ_n) 的相应子列. 因为对所有的 $x\in A$，存在 $e>0$，使 $\|x\|\geqslant e$（这 $\|\cdot\|$ 例如是R^n 上的欧几里得范数），于是对所有的 $n\in\mathbf{N}$，有

$$|\mu_n| = \frac{|\mu_n|\,\|b_n\|}{\|b_n\|}\leqslant\frac{1}{e}\|\mu_n b_n\|.$$

从而由序列 $(\mu_n b_n)$ 的有界性推出 (μ_n) 的有界性，因此，存在 (μ_n) 的子列 (ν_n)及某个$\nu\geqslant 0$，使$\nu_n\to\nu$. 设(c_n)是与 (b_n) 相应的子列，则

$$\nu_n c_n\to x,\ \nu_n\to\nu,\ c_n\to b\ (n\to\infty)$$

因此 $x=\nu b$，这就证得所述结果.

7.8 定理

设 g 是 R^n 上的实 凸 函数且满足 Slater 条件. $C=\{x\in R^n \mid g(x)\leqslant 0\}$，假定 $x_0\in R^n$，使得 $g(x_0)=0$. 则 $\partial\delta_C(x_0)$ 是由$\partial g(x_0)$ 生成的锥.

证明. 设 h是定理7.6中定义的函数，K是定理7.5中定义的锥. 由定理7.5有

$$\partial\delta_C(x_0)=K^0=\overline{K^0}=\{x\in\mathbf{R}^n\mid g'(x_0;\ x)\leqslant 0\}^{00} \quad (55)$$

（由于 h 的连续性）。根据定理7.6，$h=\delta_D^*$ 这里 $D=\partial g(x_0)$，从而

$$\begin{aligned}\{x\in\mathbf{R}^n\mid g'(x_0;\ x)\leqslant 0\}&=\{x\in\mathbf{R}^n\mid(\forall y\in D)(x\mid y)\leqslant 0\}\\&=\{x\in\mathbf{R}^n\mid(\forall y\in K_D)(x\mid y)\leqslant 0\}\\&=K_D^0\end{aligned}$$

这里 K_D 是由 D 生成的锥，即

$$K_D=\{\lambda x\mid\lambda\geqslant 0,\ x\in D\}.$$

因为 D 是凸的，K_D 是一个凸锥，由§3.13，$K_D^{00}=\overline{K_D}$，因此（55）式导出

$$\partial\delta_C(x_0)=K_D^{00}=\overline{K_D}. \quad (56)$$

读者易证 $\partial g(x_0)$ 是闭的。因为 h 仅取有限值，$\partial g(x_0)$ 就有界（见第六章练习10）。故 $\partial g(x_0)$ 是紧集。根据 Slater 条件，$0\notin\partial g(x_0)$。应用引理7.7得知 K_D 是闭集，因此由（56）式，$\partial\delta_C(x_0)=K_D$。

7.9

$\mathbf{R}^n$ 中的凸规划问题是寻求实凸函数 f 在受制于 p 个约束

$$g_i(x)\leqslant 0 \qquad (1\leqslant i\leqslant p) \quad (57)$$

时的极小值，这里 g_1，g_2，…，g_p 是 $\mathbf{R}^n$ 上的实凸函数。简记为

$$\begin{cases}\min & f(x)\\ g_i(x)\leqslant 0 & (1\leqslant i\leqslant p).\end{cases} \quad \text{(CP)}$$

满足 $g_i(x_0)\leqslant 0$ $(1\leqslant i\leqslant p)$ 的点 $x_0\in\mathbf{R}^n$ 称为（CP）的可行解。满足 $g_i(x_0)<0$ $(1\leqslant i\leqslant p)$ 的点 $x_0\in\mathbf{R}^n$ 称为（CP）的严

格可行解。点$x_0 \in \mathbf{R}^n$称为（CP）的最优解，若它是（CP)的可行解且满足

$$f(x_0) = \min\{f(x) \mid g_i(x) \leqslant 0 \quad (1 \leqslant i \leqslant p)\}.$$

（CP）称为是线性规划，如果f是线性函数，g_1，g_2，…，g_n是仿射函数。对于线性规划（57）式可以写成

$$Ax \leqslant b$$

这里A：$\mathbf{R}^n \to \mathbf{R}^p$是线性映射，$b \in \mathbf{R}^p$。

定理. 假设凸规划问题（CP）有一个严格可行解。则点$x_0 \in \mathbf{R}^n$是（CP）的最优解当且仅当存在$x' \in \partial f(x_0)$及λ_1，λ_2，…，$\lambda_p \in \mathbf{R}$，使得

$$\begin{cases} -x' \in \lambda_1 \partial g_1(x_0) + \lambda_2 \partial g_2(x_0) + \cdots + \lambda_p \partial g_p(x_0) \\ g_i(x_0) \leqslant 0, \ \lambda_i \geqslant 0, \ \lambda_i g_i(x_0) = 0 \quad (1 \leqslant i \leqslant p) \end{cases} \tag{58}$$

证明. 定义$C_i := \{x \in \mathbf{R}^n \mid g_i(x) \leqslant 0\}$ $(1 \leqslant i \leqslant p)$ 和

$$C := \bigcap_{i=1}^{p} C_i.$$

由§7.2，点$x_0 \in \mathbf{R}^n$是（CP）的最优解，当且仅当存在$x' \in \partial f(x_0)$，使得$-x' \in \partial \delta_C(x_0)$。我们有

$$\delta_C = \sum_{i=1}^{p} \delta_{C_i}.$$

由假设推出

$$\bigcap_{i=1}^{p} \operatorname{int}(C_i) \neq \phi.$$

应用定理5.38，我们可以重述条件$-x' \in \partial \delta_C(x_0)$为

$$-x' \in \partial\delta_{C_1}(x_0) + \partial\delta_{C_2}(x_0) + \cdots + \partial\delta_{C_p}(x_0). \qquad (59)$$

由§7.3，$\partial\delta_{C_i}(x_0) = \{0\}$，若$g_i(x_0) < 0$（这等价于$x_0 \in \text{int}(C_i)$），又由定理7.8，$\partial\delta_{C_i}(x_0)$ 是由$\partial g_i(x_0)$ 生成的锥，若$g_i(x_0) = 0$。从而，(59) 式等价于 $x_0 \in C$及存在 $\lambda_1, \lambda_2, \cdots, \lambda_p \geqslant 0$，使得

$$-x' \in \lambda_1 \partial g_1(x_0) + \lambda_2 \partial g_2(x_0) + \cdots + \lambda_p \partial g_p(x_0)$$

且若$g_i(x_0) < 0$ 则$\lambda_i = 0$ 。

读者应注意（58）式中的第一个条件可以重述如下：

函数$f + \sum_{i=1}^{p} \lambda_i g_i$在$x_0$有极小。

如果函数 $f, g_1, \cdots, g_p$ 是 G—可微的，定理 5.37 推出（58）式可以写成

$$\begin{cases} \nabla f(x_0) + \lambda_1 \nabla g_1(x_0) + \lambda_2 \nabla g_2(x_0) + \cdots + \lambda_p \nabla g_p(x_0) = 0 \\ g_i(x_0) \leqslant 0, \ \lambda_i \geqslant 0, \ \lambda_i g_i(x_0) = 0 \quad (1 \leqslant i \leqslant p) \end{cases} \qquad \text{(KT)}$$

条件（KT）叫做*Kuhn—Tucker*条件。这些λ_i叫做(Lagrange)乘子，它们起着类于初等微积分中拉格朗日乘子所起的作用，在那里是极小化一个受制于等式约束的函数。读者注意，这里是不等式约束，而乘子是非负的。

7.10 例

极小化函数$x + y$，使满足约束

$$\begin{cases} x^2 + y^2 \leqslant 1 \\ (x-1)^2 + y^2 \leqslant 1 . \end{cases}$$

读者易证在我们问题中的所有 函 数 都 是 凸 的（见定理5.29）．*Kuhn—Tucker*条件是：

$$\begin{cases} 1+2\lambda_1 x+2\lambda_2(x-1)=0 \\ 1+2\lambda_1 y+2\lambda_2 y=0 \\ x^2+y^2\leqslant 1 \\ (x-1)^2+y^2\leqslant 1 \\ \lambda_1\geqslant 0,\ \lambda_2\geqslant 0. \\ \lambda_1(x^2+y^2-1)=0 \\ \lambda_1[(x-1)^2+y^2-1]=0. \end{cases}$$

由此得到 $\lambda_1=0$， $\lambda_2=\frac{1}{2}\sqrt{2}$， $x=1-\frac{1}{2}\sqrt{2}$， $y=-\frac{1}{2}\sqrt{2}$。于是满足给定约束的函数 $x+y$ 在点（$1-\frac{1}{2}\sqrt{2}$，$-\frac{1}{2}\sqrt{2}$）达到极小，极小值是 $1-\sqrt{2}$。

鞍　　点

7.11

再次考虑凸规划问题

$$\begin{cases} \min f(x) \\ g_i(x)\leqslant 0 \quad (1\leqslant i\leqslant p) \end{cases} \qquad \text{(CP)}$$

这里 f，g_1，g_2，…，g_p 是 $\mathbf{R}^n$ 上的实凸函数。下面我们将借助于鞍点来刻划（CP）的最优解。

定义

(a)　我们定义函数 L：$\mathbf{R}^n\times R^p\longrightarrow\mathbf{R}$ 为

$$L(x, y)=f(x)+\sum_{i=1}^{p}\eta_i g_i(x)$$

$(x\in \mathrm{R}^n,\ y=(\eta_1,\ \eta_2,\ \cdots,\ \eta_p)\in \mathrm{R}^p)$。$L$ 称为问题（CP）的拉格朗日算子（或拉格朗日函数）。

（b） 用P^p表示非负卦限$\{x\in \mathrm{R}^p \mid x\geqslant 0\}$。$(x_0,\ \lambda)\in \mathrm{R}^n\times P^p$称为是$L$在 $\mathrm{R}^n\times P^p$上的鞍点，如果对所有的$(x,\ y)\in \mathrm{R}^n\times P^p$有

$$L(x_0,\ y)\leqslant L(x_0,\ \lambda)\leqslant L(x,\ \lambda)。$$

7.12 定理

假设凸规划问题（CP）有严格可行解。则点 $x_0\in R^n$ 是（CP）的最优解，当且仅当存在 $\lambda\in P^p$，使得 $(x_0,\ \lambda)$ 是L在$\mathrm{R}^n\times P^p$上的鞍点。

证明。充分性：$L(x_0,\ y)\leqslant L(x_0,\ \lambda)$ 蕴涵

$$\sum_{i=1}^{p}(\eta_i-\lambda_i)g_i(x_0)\leqslant 0。\qquad (60)$$

因为对所有的$y\in P^p$（60）式成立，我们有

$$g_i(x_0)\leqslant 0\qquad (1\leqslant i\leqslant p)。\qquad (61)$$

若选取$y=0$，我们得$L(x_0,\ 0)\leqslant L(x_0,\ \lambda)$，因此

$$\sum_{i=1}^{p}\lambda_i g_i(x_0)\geqslant 0。$$

结合最后这个不等式与（61）式便得出

$$\sum_{i=1}^{p}\lambda_i g_i(x_0)=0$$

从而不等式

$$L(x_0, \lambda) \leqslant L(x, \lambda) \qquad (x \in \mathrm{R}^n)$$

变成

$$f(x_0) \leqslant f(x) + \sum_{i=1}^{p} \lambda_i g_i(x) \quad (x \in \mathrm{R}^n). \qquad (62)$$

由（62）式推出，对所有 $x \in \mathrm{R}^n$，$g_i(x) \leqslant 0$（$1 \leqslant i \leqslant p$）有 $f(x) \geqslant f(x_0)$。即x_0是（CP）的最优解。

必要性。由§7.9，若 $x_0 \in \mathrm{R}^n$ 是（CP）的最优解，则存在$\lambda \in P^p$，使$\lambda_i g_i(x_0) = 0$（$1 \leqslant i \leqslant p$），且函数

$$f + \sum_{i=1}^{p} \lambda_i g_i$$

在x_0有极小。由此得出

$$L(x_0, \lambda) \leqslant L(x, \lambda) \qquad (x \in \mathrm{R}^n). \qquad (63)$$

因为$g_i(x_0) \leqslant 0$（$1 \leqslant i \leqslant p$），对每个$y = (\eta_1, \eta_2, \cdots, \eta_p) \in P^p$有

$$\sum_{i=1}^{p} \eta_i g_i(x_0) \leqslant 0$$

因此

$$L(x_0, y) = f(x_0) + \sum_{i=1}^{p} \eta_i g_i(x_0) \leqslant f(x_0)$$

$$= f(x_0) + \sum_{i=1}^{p} \lambda_i g_i(x_0) = L(x_0, \lambda).$$

结合最后结果与（63）式，便得知（x_0，λ）是L在 $\mathrm{R}^n \times P^p$ 上的鞍点。

注。 在证明的第一部分中，我们既没有用到给定函数的凸性，也没有用到（CP）的严格可行解的存在性。

7.13

我们不用§7.9的结果，可给出定理7.12的必要性的一个替换证明．定义 A 为所有这样的 $(\alpha, y)\in \mathbf{R}\times\mathbf{R}^p$ 组成的集合：对某个 $x\in\mathbf{R}^n$

$$f(x)-f(x_0)<\alpha \text{ 且 } g_i(x)\leqslant\eta_i \quad (1\leqslant i\leqslant p)$$

（这里 $y=(\eta_1, \eta_2, \cdots, \eta_p)$）．$A$ 是凸集，且 $(0, 0)\notin A$．由§4.11，在 $\mathbf{R}\times\mathbf{R}^p$ 中存在一个超平面真分离 $(0, 0)$ 和 A．从而存在 $(\beta, u)\in R\times R^p$，$(\beta, u)\neq(0, 0)$，使得对所有的 $(\alpha, y)\in A$ 有

$$\beta\alpha+(y|u)\geqslant 0. \qquad (64)$$

令 $\alpha\to+\infty$，$\eta_i\to+\infty$ $(i=1,2,\cdots,p)$，就得 $\beta\geqslant 0$ 和 $u\in P^p$．在（64）式中，令 $\eta_i=g_i(x)$ $(1\leqslant i\leqslant p)$ 且令 $\alpha\downarrow f(x)-f(x_0)$，我们得

$$\beta[f(x)-f(x_0)]+\sum_{i=1}^{p}u_i g_i(x)\geqslant 0 \qquad (x\in\mathbf{R}^n). \qquad (65)$$

假若 $\beta=0$．那么由于严格可行解的存在，（65）式蕴涵 $u=0$，矛盾．从而 $\beta>0$．用 β 除（65）式，则对某个 $\lambda=(\lambda_1,\lambda_2\cdots,\lambda_p)\in P^p$ 有

$$f(x_0)\leqslant f(x)+\sum_{i=1}^{p}\lambda_i g_i(x) \qquad (x\in\mathbf{R}^n). \qquad (66)$$

在（66）式中令 $x=x_0$，即得

$$\sum_{i=1}^{p}\lambda_i g_i(x_0)\geqslant 0.$$

所以$\sum_{i=1}^{p}\lambda_i g_i(x_0)=0$。这样（66）式推出

$$L(x_0,\lambda)\leqslant L(x,\lambda)\quad (x\in\mathbf{R}^n).$$

再循着定理7.12中从（63）式开始后面的证明，即可导出结论。

Fenchel 对偶定理

7.14

在凸规划中，极小化问题常常转换成某个确定的极大化问题，它被称为对偶问题。有不同的构造对偶问题的方法；作为一个例子，我们证明Fenchel对偶定理。

设E是R上的线性赋范空间（含多于一点），用$x\to\|x\|$表示范数，E'表示对偶。函数g：$E\to\overline{\mathbf{R}}$称为凹的，如果$-g$是凸的。对凹函数$g$，我们定义

$$\mathrm{dom}(g)=\{x\in E \mid g(x)>-\infty\}.$$

g称为是正常凹的，如果$-g$是正常凸的，凹函数g的共轭g^*：$E'\to\overline{\mathbf{R}}$定义成

$$g^*(x')=\inf_{x\in E}\{\langle x|x'\rangle-g(x)\}\quad (x'\in E').$$

g^*是凹函数，且对所有的$x'\in E'$，

$$g^*(x')=-(-g)^*(-x').$$

7.15 定理（Fenchel对偶定理）

设f是E上的正常凸函数，g是E上的正常凹函数，则在下列每种情形中：

(a) 存在点属于 $\mathrm{dom}(f)\cap\mathrm{dom}(g)$，使得 f 或 g 在此点连续；

(b) $E=\mathrm{R}^n$ 且 $\mathrm{ri}(\mathrm{dom}(f))\cap\mathrm{ri}(\mathrm{dom}(g))\neq\phi$。我们有

$$\inf_{x\in E}\{f(x)-g(x)\}=\max_{x'\in E'}\{g^*(x')-f^*(x')\}.$$

证明。对所有的 $x\in E$，$x'\in E'$ 我们有

$$f(x)+f^*(x')\geqslant\langle x|x'\rangle\geqslant g(x)+g^*(x')$$

因此

$$f(x)-g(x)\geqslant g^*(x')-f^*(x')$$

于是

$$m:=\inf_{x\in E}\{f(x)-g(x)\}\geqslant\sup_{x'\in E'}\{g^*(x')-f^*(x')\}. \tag{67}$$

若 $m=-\infty$，所述结果显然成立。现在假定 $m>-\infty$。由假设推出 $m\in\mathrm{R}$。定义

$$C_1:=\{(x,\lambda)\in E\oplus\mathrm{R}|\ f(x)\leqslant\lambda\}$$

$$C_2:=\{(x,\lambda)\in E\oplus\mathrm{R}|\ \lambda\leqslant g(x)+m\}.$$

C_1 和 C_2 是 $E\oplus\mathrm{R}$ 的凸子集。读者易证（见 § 5.38和 § 5.39）在 $E\oplus\mathrm{R}$ 中存在一个非垂直的闭超平面 $H=F^{-1}(\beta)$ 真分离 C_1 和 C_2。假定 $F(C_1)\geqslant\beta\geqslant F(C_2)$，$F=(x',\alpha)$，这里 $x'\in E'$，$\alpha\in\mathrm{R}$。我们有 $\alpha>0$，且对每个 $x\in\mathrm{dom}(f)\cap\mathrm{dom}(g)$ 成立

$$\langle x|x'\rangle+\alpha f(x)\geqslant\beta\geqslant\langle x|x'\rangle+\alpha\{g(x)+m\}$$

因此

$$f(x)\geqslant\langle x|y'\rangle-\gamma\geqslant g(x)+m$$

这里 $y' = -x'/\alpha \in E'$，$\gamma = -\beta/\alpha$。此外，容易看出对所有的 $x \in E$，这些不等式仍成立。从而

$$f^*(y') \leqslant \gamma \text{ 和 } g^*(y') \geqslant \gamma + m$$

于是由（67）式得

$$m \leqslant g^*(y') - f^*(y') \leqslant \sup_{x' \in E'} \{g^*(x') - f^*(x')\} \leqslant m$$

这样证明了所述结果。

7.16 例

设 $C \subset E$ 是非空凸集，$x_0 \in E$。定义

$$f(x) := \|x - x_0\|,\quad g(x) := -\delta_C(x) \quad (x \in E)$$

应用 Fenchel 对偶定理得出

$$\inf_{x \in C} \|x - x_0\| = \inf_{x \in E} \{f(x) - g(x)\}$$

$$= \max_{x' \in E'} \{g^*(x') - f^*(x')\}$$

从例6.7和§6.3的性质（g）得

$$f^*(x') = \delta_S(x') + \langle x_0 | x' \rangle \quad (x' \in E')$$

这里 $S = \{x' \in E' \mid \|x\| \leqslant 1\}$。我们有 $g^*(x') = -\delta_C^*(-x')$

从而

$$\inf_{x \in C} \|x - x_0\| = \max_{x' \in E'} \{-\delta_C^*(-x') - \delta_S(x') - \langle x_0 | x' \rangle\}$$

$$= \max_{x' \in S} \{\langle -x_0 | -x' \rangle - \delta_C^*(-x')\}$$

$$= \max_{\|x\| \leqslant 1} \{\langle -x_0 | x' \rangle - \delta_C^*(x')\}。$$

当 $E = R^n$ 时，读者给出这个公式的几何解释。

邻 近 映 射

7.17

设 f 是 R^n 上的下半连续正常凸函数. $x_0\in\mathrm{R}^n$. 定义函数 f: $\mathrm{R}^n\to\mathrm{R}$为

$$F(x)=f(x)+\frac{1}{2}\|x-x_0\|^2\quad(x\in R^n)$$

这里 $x\to\|x\|$是 R^n 上的欧几里得范数. F 是下半连续凸函数的证明留给读者。我们将证F有整体极小。由引理6.12，f 有连续仿射弱函数，于是存在 $a\in\mathrm{R}^n$，$\alpha\in\mathrm{R}$，对所有的 $x\in\mathrm{R}^n$满足

$$f(x)\geqslant(x|a)-\alpha$$

从而，对每个 $x\in\mathrm{R}^n$

$$F(x)\geqslant(x|a)-\alpha+\frac{1}{2}\|x-x_0\|^2$$

$$=\frac{1}{2}\|x-(x_0-a)\|^2+(x_0|a)-\frac{1}{2}\|a\|^2-\alpha.$$

设 $b\in\mathrm{dom}(f)$。最后这个不等式可推出存在数 $R>0$，使得对任何的$\|x-(x_0-a)\|>R$有$F(x)>F(b)$。设F在紧球$\{x|\ \|x-(x_0-a)\|\leqslant R\}$上的限制有极小值为$m$（见§5.4性质(b)）。假定这极小值在点 c 取得。因为$m\leqslant F(b)$，从而F在 c 有整体极小。假设 $d\in\mathrm{R}^n$ 满足$F(d)=m$。因为

$$\left\|\frac{1}{2}(c+d)-x_0\right\|^2=\frac{1}{2}\|c-x_0\|^2+\frac{1}{2}\|d-x_0\|^2-\frac{1}{4}\|c-d\|^2$$

我们有

$$m \leqslant F\left(\frac{1}{2}(c+d)\right)$$

$$= f\left(\frac{1}{2}(c+d)\right) + \frac{1}{2}\left\|\frac{1}{2}(c+d) - x_0\right\|^2$$

$$\leqslant \frac{1}{2}f(c) + \frac{1}{2}f(d) + \frac{1}{4}\|c - x_0\|^2 + \frac{1}{4}\|d - x_0\|^2$$

$$-\frac{1}{8}\|c-d\|^2$$

$$\leqslant m - \frac{1}{8}\|c-d\|^2$$

因此 $c=d$。这样，F的极小点是唯一的。这个点记成

$$\mathrm{prox}_f(x_0)。$$

从R^n到它自身的映射prox_f叫做（关于f的）邻近映射。

注。

(a) 利用F的严格凸性也可证明极小点的唯一性（§5.18 例（a）和第五章练习8）。

(b) 设$C \subset \mathrm{R}^n$是非空闭凸集，$f = \delta_C$（见定理6.20），则$\mathrm{prox}_f(x_0)$是C中对x_0的最佳逼近，即是C中最接近x_0的点（也叫做x_0在C上的投影）。

7.18 定理

设f是R^n上的下半连续正常凸函数，x，y，$z \in \mathrm{R}^n$。则下列条件是等价的：

(a) $z = x + y$和$f(x) + f^*(y) = (x|y)$。

(b) $x = \mathrm{prox}_f(z)$及$y = \mathrm{prox}_{f^*}(z)$。

证明，定义从 R^n 到 R 的函数 g 和 F 为 $g(u)=\frac{1}{2}\|u-z\|^2$，$F=f+g$，由定理5.38有 $\partial F=\partial f+\partial g$。从§7.1得出

$$x=\mathrm{prox}_f(z)\Longleftrightarrow 0\in\partial F(x)\Longleftrightarrow 0\in\partial f(x)+\partial g(x).$$

函数 g 是 Fréchet 可微的，且对每个 $u\in R^n$，$\nabla g(u)=u-z$。定理5.37推出

$$x=\mathrm{prox}_f(z)\Longleftrightarrow 0\in\partial f(x)+x-z$$

$$\Longleftrightarrow(\text{存在}y\in\partial f(x))z=x+y \quad (68)$$

及

$$y=\mathrm{prox}_{f^*}(z)\Longleftrightarrow(\text{存在}x\in\partial f^*(y))z=x+y. \quad (69)$$

定理6.10和6.18蕴涵

$$y\in\partial f(x)\Longleftrightarrow f(x)+f^*(y)=(x|\,y)\Longleftrightarrow x\in\partial f^*(y) \quad (70)$$

(a)$\Longrightarrow$(b)：结合 (a) 和 (70) 式得到 $y\in\partial f(x)$ 及 $x\in\partial f^*(y)$，由此及 (a)、(68)和 (69) 式便得 $x=\mathrm{prox}_f(z)$，$y=\mathrm{prox}_{f^*}(z)$。

(b)$\Longrightarrow$(a)：由 (b) 和 (68) 式，存在点 $y_0\in\partial f(x)$，使 $z=x+y_0$。(70) 式蕴涵 $x\in\partial f^*(y_0)$ 及 $f(x)+f^*(y_0)=(x|\,y_0)$。由式及 (69) 式得 $y_0=\mathrm{prox}_{f^*}(z)=y$。从而 $z=x+y$ 且 $f(x)+f^*(y)=(x|\,y)$。

7.19 例

设 $K\subset R^n$ 是非空闭凸锥。令 $f=\delta_K$。读者易证 $f^*=\delta_{K^\circ}$。条件

$$\delta_K(x)+\delta_{K^\circ}(y)=(x|\,y)$$

等价于

$$x \in K, \quad y \in K^{\circ}, \quad (x|y) = 0.$$

这样，公式

$$z = \text{prox}_f(z) + \text{prox}_{f^*}(z)$$

给出 z 的唯一正交分解作为分别是K和K^0中元素的和（即z在K和 K^0 上的投影）。

单 调 算 子

7.20

设E 是R上的线性赋范空间。定义从E到E' 的多值函数是E到E'的幂集$P(E')$ 的映射。多值函数 T: $E \to E'$ 称为单调算子，如果对任何的 x，$y \in E$，x'，$y' \in E'$，$x' \in Tx$，$y' \in Ty$，有

$$\langle x-y | x'-y' \rangle \geqslant 0.$$

注意从R到 R的非减函数是单调算子。

多值函数T: $E \to E'$称为是循环单调算子，如果对任意有限个对 (x_1, x_1')，(x_2, x_2')，…，(x_p, x_p')，其中 $x_i \in E$，$x_i' \in Tx_i$ $(1 \leqslant i \leqslant p)$，有

$$\langle x_1 - x_2 | x_1' \rangle + \langle x_2 - x_3 | x_2' \rangle + \cdots + \langle x_{p-1} - x_p | x_{p-1}' \rangle + \langle x_p - x_1 | x_p' \rangle \geqslant 0 \qquad (71)$$

在 (71) 式中令$p = 2$，我们看到循环单调算子在特殊情况下是单调算子，但其逆不真。

单调算子在优化理论中起着重要作用。下述的定理指出单调算子与凸函数间的某些联系。

7.21 定理

设 f 是E上的正常凸函数。则∂f是循环单调算子。

证明。设$x_i\in E$，$x'_i\in\partial f(x_i)$ $(1\leqslant i\leqslant p)$ 。则

$$f(x_2)\geqslant f(x_1)+\langle x_2-x_1|x'_1\rangle$$
$$f(x_3)\geqslant f(x_2)+\langle x_3-x_2|x'_2\rangle$$
$$\vdots \qquad \vdots \qquad \vdots$$
$$f(x_p)\geqslant f(x_{p-1})+\langle x_p-x_{p-1}|x'_{p-1}\rangle$$
$$f(x_1)\geqslant f(x_p)+\langle x_1-x_p|x'_p\rangle.$$

将这些不等式相加便得

$$0\geqslant\langle x_2-x_1|x'_1\rangle+\langle x_3-x_2|x'_2\rangle+\cdots+\langle x_1-x_p|x'_p\rangle$$

所述结果得证。

7.22

一个（循环）单调算子T：$E\to E'$称为是最大（循环）单调，如果它的图形

$$\{(x,\ x')\in E\times E'|\ x'\in Tx\}$$

不全含于其它任何从E到E'的（循环）单调算子的图形中。

定理。设 f 是 $\mathbb{R}^n$ 上的下半连续正常凸函数，则 ∂f是最大单调和最大循环单调。

证明。假设x_0，$y_0\in\mathbb{R}^n$，对所有x，$y\in\mathbb{R}^n$，且 $y\in\partial f(x)$，成立

$$(x-x_0|\ y-y_0)\geqslant 0. \tag{72}$$

令

$$x_1=\mathrm{prox}_f(x_0+y_0),\quad y_1=\mathrm{prox}_{f^*}(x_0+y_0)$$

（见§7.17）。定理7.18蕴涵$x_0+y_0=x_1+y_1$及 $y_1\in\partial f(x_1)$。

在（72）式中令$x=x_1$，$y=y_1$得到

$$-\|x_0-x_1\|^2=(x_1-x_0|x_0-x_1)\geqslant 0$$

（这里$x\to\|x\|$是R^n上的欧几里得范数）因此$x_0=x_1$，于是$y_0=y_1\in\partial f(x_1)=\partial f(x_0)$。我们得出$\partial f$是最大单调。现在从定理7.21得出$\partial f$也是最大循环单调。

注 释

1. 许多凸规划问题来源于数理经济学。例如，见

H. Nikaido, *Convex Structures and Economic Theory*, New York, Academic Press, 1968.

在含有不等式约束的最优化问题中，强调凸性作用的最初文章之一是

H. W. Kuhn 和 A. W. Tucker, Nonlinear programming, *in Proc. and Berkeley Symp. on Mathematical Statistics and Probability*, *Berkeley* (1951) 481—92.

2. 在§7.13，为了证明（65）式中$f(x)-f(x_0)$的系数β不为0，我们利用严格可行解的存在。确保$\beta>0$的这种条件称为约束品性。其它的约束品性可在下面文献中找到：

O. L. Mangasarian, *Nonlinear Programming*, New York, Mc Graw-Hill, 1969,

M. S. Bazaraa and C. M. Shetty, *Foundations of Optimization*, Berlin, Springer (*Lecture Notes in Economics and Mathematical Systems no.* 122, 1976).

3. Fenchel对偶定理（见§7.15）的原始（有限维）形式见

W. Fenchel, *Convex Cones, Set and Functions*, Lecture notes, Princeton, 1953.

扩充到无穷维情形见

R. T. Rockafellar, Extension of Fenchel's duality theorem for convex functions, *Duke Math. J.* 33 (1966) 81-9.

4. 在下述文献中，邻近映射的理论已被发展。

J. J. Moreau, Proximit'e et dualit'e dans un espace hilbertien, *Bull. Soc. Math. France* 93 (1965) 273—99.

5. 与凸分析相联系研究单调算子的最初文章之一是：

G. J. Minty, Monotone (nonlinear) operators in Hilbert space, *Duke Math.* J. 29 (1962) 341—6

最大单调算子理论的阐述在下述文献中可以找到。

H. Brezis, *Operateurs Maximaux Monotones et Semigroupes de Contractions dans les Espaces de Hilbert*, Amsterdam, North-Holland, 1973.

单调算子理论已证明是研究椭圆型非线性偏微分方程的有力工具。许多物理系统也用单调算子来描述。例如见

V. Dolezal, *Monotone Operators and Applications in Control and Network Theory*, Amsterdam, Elsevier, 1979.

6. 在下面列举的书中，更广泛地研究了最优化理论。

V. Barbu and Th. Precupanu, *Convexity and Optimizatiou in Banach Spaces*, Alphen aan den Rijn, The Netherlands, Sythoff & Noordhoff, 1978.

I. Ekeland and R. Temam, *Convex Analysis and Varia-*

tional Problems, Amsterdam, North-Holland, 1976.

I. V. Girsanov, *Lectures on Mathematical Theory of Extremum Problems*, Berlin, Springer (*Lecture Notes in Economics and Mathematical Systems no.* 67, 1972).

A. D. Ioffe and V. M. Tihomirov, *Theory of Extremal Problems*, Amsterdam, North-Holland, 1979.

P. J. Laurent, *Approximation et Optimisation*, Paris, Hermann, 1972.

R. Wets, *Grundlagen Konvexer Optimierung*, Berlin, Springer (*Lecture Notes in Economics and Mathematical Systems no.* 137, 1976).

答案和提示

第一章

1．(a) 由定理1.6，f'_+ 存在且是非减函数．利用不等式

$$f(x) \geqslant f(y) + f'_+(y)(x-y) \quad (x,\ y \in \langle a,\ b\rangle).$$

(b) 设 x 是 f 的局部极小点，及 $y \in \langle a,\ b\rangle$。考虑点 $z = \lambda x + (1-\lambda)y$，这里 $\lambda \in \langle 0,\ 1\rangle$。如果 λ 充分接近 1，$f(z) \geqslant f(x)$。从而 $f(y) \geqslant f(x)$。

(c) 假设 x 和 y 是 f 的整体极小点．考虑点 $\frac{1}{2}(x+y)$．

2．f 在 c 可微，当且仅当 $f'_+(c) = f'(c)$．

3．至少存在一条过 $(x, f(x))$ 的直线处处不位于 f 图形之上方，当且仅当存在 $m \in \mathrm{R}$，使得对任何的 $y \in \langle a,\ b\rangle$，

$$f(y) \geqslant f(x) + m(y-x).$$

充分性：设 $a<x<y<b$，$0<\lambda<1$。令 $z := \lambda x + (1-\lambda)y$，考虑过 $(z,\ f(z))$ 且处处不位于 f 图形上方的直线．

4．(a) 考虑 $\exp\left(\sum_{i=1}^{n} r_i \log x_i\right)$．

(b) 在 (a) 中取$r_i=1/n$ ($1\leqslant i\leqslant n$)。

(c) 对每个积

$$\left[\frac{a_i^p}{\sum_{i=1}^{n}a_i^p}\right]^{1/p}\cdot\left[\frac{b_i^q}{\sum_{i=1}^{n}b_i^q}\right]^{1/q}\quad(1\leqslant i\leqslant n)$$

应用 (a)。

5. 设$x_0\in\mathrm{R}$。对所有的$x\in\mathrm{R}$，我们有

$$f(x)\geqslant f(x_0)+f'_+(x_0)(x-x_0)。$$

假定$f'_+(x_0)>0$。则$\lim\limits_{x\to+\infty}f(x)=+\infty$，矛盾。

类似地，假设$f'_+(x_0)<0$也导致矛盾。从而$f'_+(x_0)=0$。同理可得$f'_-(x_0)=0$。

6. (a)，(b) 应用定理1.11。

(c) 注意，若$a>0$，条件$ab\geqslant c^2$等价于函数$ax^2+2cx+b$半正定；取$a=f$，$b=f''$，$c=f'$。

(d) 应用 (c)。

7. 从定义1.1 (a) 直接可得。

8. 设$y_0\in f$ ([0，∞>)。令$H(x):=xy_0-F(x)$ ($x>0$)，则$H'(x)=y_0-f(x)$。从而$H'(x_0)=0$，若x_0满足$f(x_0)=y_0$ (因此$x_0=g(y_0)$)。

9. 必要性：如果$f(a)=+\infty$，或$f(b)=+\infty$，则对每个$\lambda\in\langle0,1\rangle$，$\lambda f(a)+(1-\lambda)f(b)=+\infty$。

充分性：利用定义1.19。

10. (a) 从§1.25定义 (a) 直接可得。

(b) 考虑

$$f(x)=\begin{cases}1 & 若 x=0\\ 0 & 若 x\neq 0.\end{cases}$$

(c) 假设$a<x<z<y<b$且$f(x)\leqslant f(y)$，$f(z)>f(y)$。鉴于f的连续性，存在$t\in\langle z, y\rangle$使得$f(t)>f(y)$。现在考虑区间$\langle x, t\rangle$。

第二章

1. 利用§2.4性质（d）。

2. 应用定理2.3。

3. 设$d(a, C)\leqslant \varepsilon$，$d(b, C)\leqslant \varepsilon$及$\lambda\in\langle 0, 1\rangle$。对每个$\delta>0$，存在$c, d\in C$满足$d(a, c)<\varepsilon+\delta$，$d(b, d)<\varepsilon+\delta$。从而

$$d(\lambda a+(1-\lambda)b, \lambda c+(1-\lambda)d)<\varepsilon+\delta,$$

因此 $d(\lambda a+(1-\lambda)b, C)\leqslant\varepsilon$。

4. 假设$C=\mathrm{co}(a_1, a_2, \cdots, a_n)=\mathrm{co}(b_1, b_2, \cdots, b_m)$，且任意$a_i(b_i)$都不是其余$a_k(b_k)$的凸组合。设$1\leqslant i\leqslant m$。表示每个$b_i$为凸组合。

$$\sum_{k=1}^{n}\lambda_k a_k$$

及每个a_k为凸组合

$$\sum_{l=1}^{m}\mu_{kl}b_l.$$

推出对某个k有$b_i=a_k$。

5. 注意$C=\mathrm{co}(e_1, e_2, \cdots, e_n)$，这里$e_1, e_2, \cdots, e_n$是$\mathbf{R}^n$中单位向量（$e_1=(1,0,0,\cdots,0)$，$e_2=(0,1,0,\cdots,0)$，

等等）。这些向量是C的顶点。

6. 设C是凸多胞形。C的每个顶点a是C的一个极端点：设C的顶点是a，x_1，x_2，…，x_k。假设存在x，$y\in C$使$a\in\langle x, y\rangle$。把x和y表示为a，x_1，x_2，…，x_k的凸组合，导出$a=x=y$。

C的每个极端点b是C的一个顶点：设C的顶点是x_1，x_2，…，x_k。把b表为凸组合

$$\sum_{i=1}^{k}\lambda_i x_i$$

并证明对某个i，$\lambda_i=1$。

7. D是V中所有可以表示成形为

$$\sum_{i=1}^{k}\lambda_i x_i$$

的点之集合，这里$k\in\mathbf{N}$且$\lambda_k<0$。

$C^i=\phi$：令

$$x=\sum_{i=1}^{k}\lambda_i x_i\in C.$$

则对任何$e<0$，$x+ex_{k+1}\notin C$。

D的凸性直接从定义2.1（b）得出。

$D^i=\phi$：令

$$x=\sum_{i=1}^{k}\lambda_i x_i\in D.$$

则对任何$e>0$，$x+ex_{k+1}\notin D$。

$D^*=V$：令

$$z=\sum_{i=1}^{m}\mu_i x_i\in V.$$

那末$[-x_{m+1}, z\rangle \subset D$.

8. (a) A和B的凸性从定义2.1(b)直接得到. 因为 $P(0) = P(2\cdot 0) = 2P(0)$, 就有 $P(0)=0$, 因此 $0\in A\cap B$. 若 $x\in V$, 则

$$P\left(\frac{1}{P(x)+1}\cdot x\right)<1$$

从而 $0\in A^i\cap B^i$.

(b) 从 (c) 和§2.21得到

(c) 设q是A的度规, 则

$$\begin{aligned} q(x) &= \inf\{\lambda\geqslant 0 \mid x\in\lambda A\} \\ &= \inf\{\lambda\geqslant 0 \mid p(x)<\lambda\} = p(x). \end{aligned}$$

因此$q=p$, 类似地, p是B的度规.

9. 若对所有的$\lambda>0$, $p(\lambda x)\leqslant 1$, 则$p(x)=0$.

10. p是连续函数, $\{x \mid p(x)<1\}$是C的开子集, 因此C是凸体.

反之, 假设C是凸体. 证明 $0\in \mathrm{int}(C)$ (见定理2.23). 从而存在0的一个邻域U, 使得$U\subset C$, 因此$p(U)\leqslant 1$. 推出p在0连续. 现在证明p在E的每点连续. 为此, 对于一个凸体C, 有$\mathrm{int}(C)=C^i$及$\overline{C}=C^a$, 再利用$\{x \mid p(x)<1\}(=C^i)$的开性及$\{x \mid p(x)\leqslant 1\}$ $(=C^a)$ 的闭性.

11. 从定理2.27直接得到.

12. 设B是A相对于E的补. 存在x, $y\in A$, $z\in B$使$z\in\langle x, y\rangle$. 定义$\alpha=\sup\{\lambda\in\mathbb{R} \mid [z, z+\lambda(y-x)]\subset B\}$, $\beta=\sup\{\lambda\in\mathbb{R} \mid [z, z+\lambda(x-y)]\subset B\}$. 令$u=z+\alpha(y-x)$, $v=z+\beta(x-y)$ 证明$u\in A$, $v\in A$, $\langle u, v\rangle\subset B$.

13. 首先证明

$$\bigcap \overline{C_i} = \bigcap \overline{\text{int}(C_i)} \supset \overline{\bigcap \text{int}(C_i)}.$$

设 $x_0 \in \bigcap \text{int}(C_i)$。若 $x \in \bigcap \overline{C_i}$，证明 $(x, x_0] \subset \bigcap \text{int}(C_i)$。证明完成。

第 三 章

1. 由定理3.8，存在E中闭超平面$H = f^{-1}(\alpha)$ 真分离A和B。假设$f(A) \leqslant \alpha$，$f(B) \geqslant \alpha$。利用 A 和B的开性证明$f(A) < \alpha$，$f(B) > \alpha$。

2. 证明存在E的开凸子集C，D满足$A \subset C$，$B \subset D$，$C \cap D = \phi$。利用练习 1 的结果。

3. 必要性：假定$H = f^{-1}(\alpha)$是使$f(x) = \alpha$的超平面。存在$y \in E$，使得$f(y) > 0$。考虑点$x + \delta x$和$x - \delta y$，这里$\delta > 0$。

充分性：假设$x \overline{\in} \text{int}(A)$。由§3.9，存在$E$中包含$x$的超平面$H$，真分离$A$和$\{x\}$。$A$中不存在被$H$严格分离的两个点。

4. 设E是线性赋范空间，f：$E \to \mathbf{R}$ 是线性函数，使$|f|$不连续（见定理 3.7 后的注（c））。考虑集合$\{x \in E | |f(x)| \leqslant 1\}$。

5. 假设$K \neq \phi$。首先证明$0 \in \overline{K}$。导出$K + \overline{K} \supset K$。为了证明$K + \overline{K} \subset K$，设$x \in K$，$y \in \overline{K}$，证明$\frac{1}{2}(x + y) \in K$。

6. 命题直接从‘极’的定义得出。我们举出一个例子。如果$u \in (K_1 + K_2)^\circ$，则对任何$x \in K_1$，$y \in K_2$，$\lambda > 0$有

$$\langle x|u\rangle+\lambda\langle y|u\rangle=\langle x+\lambda y|u\rangle\leqslant 0.$$

分别令$\lambda\to+\infty$和$\lambda\downarrow 0$，我们得到$\langle y|u\rangle\leqslant 0$及$\langle x|u\rangle\leqslant 0$。从而有

$$(K_1+K_2)^\circ\subset K_1^\circ\cap K_2^\circ.$$

第 四 章

1. 由定理4.2，每个$x\in\text{co}(A)$可以写成A中$k+1$个点a_1，a_2，…，a_{k+1}的凸组合

$$\sum_{i=1}^{k+1}\lambda_i a_i$$

与

$$x=\sum_{i=1}^{k+1}\lambda_i a_i+0\cdot a$$

遵循定理4.2的证明，试写x为a和A中其它k个点的凸组合。

2. 应用定理4.2.

3. 利用§2.4的性质(d)和定理4.2.

4. 假设$x\notin A$. 因为$d(x,A)>0$及$A_i\to A(i\to\infty)$，存在$N\in\mathbf{N}$，使对任何$i>N$，$h(A_i,A)<\frac{1}{2}d(x,A)$，这与$x_i\to x$ $(i\to\infty)$矛盾。

5. 若A非凸，则存在x，y，$z\in\mathbf{R}^n$，使$x,y\in A$，$z\notin A$，$z\in\langle x,y\rangle$. 定义$\rho=d(z,A)$（这里d是$\mathbf{R}^n$上的欧几里得度量）。我们有$\rho>0$。若$A_i\to A$ $(i\to\infty)$，则存在$N\in\mathbf{N}$使得

$h(A, A_N) < \frac{1}{2}\rho$。找出点 x_1，$y_1 \in A_N$，$z_1 \in \langle x_1, y_1\rangle$ 满足

$d(z_1, A) \geqslant \frac{1}{2}\rho$。导出$A_N$不是凸集。

6. (a) 存在 $x \in \mathrm{ri}(C)$。若$y \in \overline{C}$，则$[x, y\rangle \subset \mathrm{ri}(C)$（见§4.8），因此 $y \in C^e$。

(b) 设E是线性拓扑空间，考虑一个在E中稠密的超平面（见定理3.7后的注）。

7. 必要性是平凡的。

充分性：利用练习6(a)的结果。

8. 充分性是平凡的。

必要性：假设$A \cap \overline{C} \neq \phi$。因为$\overline{C} = \overline{\mathrm{ri}(C)}$，我们有$A \cap \overline{\mathrm{ri}(C)} \neq \phi$。利用$A$的开性导出$A \cap \mathrm{ri}(C) \neq \phi$。

9. 设$x \in \mathrm{ri}(C)$。存在$y \in C \cap \mathrm{ri}(D)$及$z \in C$使得$x \in \langle y, z\rangle$。导出 $x \in \mathrm{ri}(D)$。

10. (a) 遵循定理4.10(c)的证明。

(b) 对有限个凸集的集族情形，遵循定理 4.10(d) 的证明。其它情况，利用区间

$$\langle 0, 1 + \frac{1}{n}\rangle \quad (n \in N)$$

构造一个反例。

11. 设x_1，x_2，…，x_{n+2}是A中$n+2$个点。它们是仿射独立的。因此存在不全为0的μ_1，μ_2，…，$\mu_{n+2} \in \mathbf{R}$，使得

$$\sum_{i=1}^{n+2} \mu_i x_i = 0 \quad 及 \quad \sum_{i=1}^{n+2} \mu_i = 0.$$

定义 $A_1 = \mathrm{co}\{x_i | \mu_i > 0\}$，$A_2 = A \backslash A_1$。见 Helly 定理的证明。

12. 由§4.11，存在 $y \in \mathrm{R}^n$，$\alpha \in \mathrm{R}$，使 $(K| y) \leqslant \alpha$ 及 $(C| y) \geqslant \alpha$。证明 $y \in K^\circ$ 和 $(C| y) \geqslant 0$。

13. K_1 和 K_2 是闭凸锥，因此 $K_1^{\circ\circ} = K_1$，$K_2^{\circ\circ} = K_2$。利用第三章练习 6 的结果。

14. 因为 AP^n 是一个锥，$(AP^n| a) < \alpha$ 蕴涵 $(AP^n| a) \leqslant 0$。后一不等式蕴涵 $A^t a \leqslant 0$。

15. 设 B 是所有 $y = (\eta_1, \eta_2, \cdots, \eta_n) \in \mathrm{R}^n$ 使 $y \geqslant 0$ 的集合，$\sum_{i=1}^{n} \eta_i = 1$。$B$ 是凸紧集，因此，AB 是凸紧集。就有：

$Ay = 0$，$y \geqslant 0$，$y \neq 0$ 在 R^n 中没有解 y。

$\Longleftrightarrow 0 \notin AB$

$\Longleftrightarrow$ 存在一个超平面，严格分离 $\{0\}$ 和 AB。

第 五 章

1，2，5，6 从定义5.9直接得出。

3. 利用§5.12。

4. (a) 从定义5.9直接得出。

(b) 利用集合 $\{x \in E | \delta_A(x) \leqslant \lambda\}$（见定理5.3）。

7. (a) 设 f 是常数。

(b) 鉴于 f 的连续性，我们有 $B \subset \mathrm{int}(A)$。设 $x \in E$，使 $f(x) < \lambda$。设 $x_0 \in \mathrm{int}(A)$。存在 $y \in A$，使得 $x_0 \in \langle x, y \rangle$。导出 $f(x_0) < \lambda$。从而有 $\mathrm{int}(A) \subset B$。

8. (a) 设 x 是 f 的局部极小点。对每个 $y \in V$，考虑 f

在过x和y的直线上的限制，证明$f(y)\geqslant f(x)$（见第一章练习1）。

(b) 假设x和y是f的整体极小点。考虑点$\frac{1}{2}(x+y)$。

9. 利用函数$x\longrightarrow\|x-x_0\|^2$的严格凸性。见§5.18的例(a)。

10. $\mathbf{dom}(f\Box g)=\mathbf{dom}(f)+\mathbf{dom}(g)$。

11. 见§5.20。如果$f(a)=-\infty$且f在a上局部有界，存在a的一个邻域U，使得$U\subset\mathbf{dom}(f)$。由§5.12，对所有的$x\in U$，$f(x)=-\infty$。

12. 应用定理5.29的类似情形。

13. 从定义5.9直接得到。

14. 设x，$y\in E$，$\lambda\in\langle 0, 1\rangle$。令$z=\lambda x+(1-\lambda)y$。设$f$: $E\to\mathbf{R}$是次可微函数，$z'\in\partial f(z)$。对所有的$t\in E$，我们有

$$f(t)\geqslant f(z)+\langle t-z|z'\rangle.$$

分别取$t=x$和$t=y$，导出$\lambda f(x)+(1-\lambda)f(y)\geqslant f(z)$。

15. 令$C=\mathrm{co}(e_1,\cdots,e_n,-e_1\cdots,-e_n)$。

证明

$$C=\left\{(\zeta_1,\zeta_2,\cdots,\zeta_n)\in\mathbf{R}^n\left|\sum_{i=1}^{n}|\zeta_i|\leqslant 1\right.\right\}$$

（见第二章练习2）。

注意$y\in\partial f(0)$等价于

$$\max_i|\zeta_i|\geqslant\langle x|y\rangle\quad\forall x=(\zeta_1,\zeta_2,\cdots\zeta_n)\in\mathbf{R}^n.\qquad(*)$$

证明(*)等价于$y\in C$。导出$\partial f(0)=C$。

16. (a) 从定义5.9直接得到。

(b) 设 $x\in E$，$x'\in E'$ 使 $\langle x|\, x'\rangle=\|x\|\cdot\|x'\|$ 且 $\|x'\|=\|x\|$。对所有的 $y\in E$，我们有

$$f(y)-f(x)-\langle y-x|\, x'\rangle$$

$$=\frac{1}{2}\|y\|^2-\frac{1}{2}\|x\|^2-\langle y|\, x'\rangle+\|x\|^2$$

$$\geqslant\frac{1}{2}\|y\|^2+\frac{1}{2}\|x\|^2-\|y\|\cdot\|x\|$$

$$=\frac{1}{2}(\|y\|-\|x\|)^2\geqslant 0。$$

反之，设 $x\in E$，$x'\in\partial f(x)$。对任何 $y\in E$ 有

$$\frac{1}{2}\|y\|^2\geqslant\frac{1}{2}\|x\|^2+\langle y-x|\, x'\rangle。$$

取 $y=\lambda x$，分别令 $\lambda\uparrow 1$ 及 $\lambda\downarrow 1$。导出 $\langle x|\, x'\rangle=\|x\|^2$，因此 $\|x\|\leqslant\|x'\|$，取 $y=x+\varepsilon z$，推出对任何 $z\in E$，$z\neq 0$

$$\frac{|\langle z|\, x'\rangle|}{\|z\|}\leqslant\|x\|$$

就有 $\|x'\|\leqslant\|x\|$。

(c) 利用 (b) 的结果。

17. 应用定理5.37并见练习14。

18. 定义从 R 到 $\overline{\mathrm{R}}$ 的函数 f_1 和 f_2 为

$$f_1(x)\begin{cases}-\sqrt{x} & \text{如果 } x\geqslant 0\\ +\infty & \text{如果 } x<0\end{cases}$$

和

$$f_2(x)=f_1(-x)$$

我们有 $\partial f_1(0)=\partial f_2(0)=\phi$ 及 $\partial(f_1+f_2)(0)=\mathrm{R}$。

19. 设 f_1 和 f_2 是 R^n 上的正常凸函数，$x\in \mathrm{dom}(f_1)\cap \mathrm{dom}(f_2)$，$f_1$在 x 连续。我们有 $x\in \mathrm{int}(\mathrm{dom}(f_1))\cap \mathrm{dom}(f_2)$。证明 $\mathrm{int}(\mathrm{dom}(f_1))\cap \mathrm{ri}(\mathrm{dom}(f_2))\neq\phi$。

20. 定义 g 为

$$g(x)=\begin{cases} f(x) & 若\ x\in A \\ +\infty & 若\ x\overline{\in} A. \end{cases}$$

我们有 $\{x\in A|\ f(x)\leqslant\lambda\}=\{x\in X|\ g(x)\leqslant\lambda\}\quad \forall\lambda\in R$。证明

(a)$\Longleftrightarrow g$是下半连续的$\Longleftrightarrow$(b)

由定理5.8，g 是下半连续的，当且仅当

$$\liminf_{x\to a} g(x)\geqslant g(a)\quad \forall a\in X \qquad (*)$$

证明 (*) 等价于 (c)。

21. (a) 用定义

$$f(x)=+\infty\quad 若\ x\overline{\in} C$$

将 f 延拓到整个 E 上。区别两种情形，对任何 $x\in\bar{C}$，$\bar{f}(x)>-\infty$，及存在 $x\in\bar{C}$，使 $\bar{f}(x)=-\infty$。证明在后一情形，对所有的 $x\in E$，我们有 $|\bar{f}(x)|=+\infty$ (见§5.12)，因此对所有的 $x\in\bar{C}$，$\bar{f}(x)=-\infty$。

应用定理5.8 (c)。

(b) 利用定理5.23和 (a) 的结果。

(c) 设 f：$E\longrightarrow R$ 是线性的，但不连续。取 $C=E$。

22. 设$A\subset U$是紧集。存在 $\varepsilon>0$，使 $A_\varepsilon\subset U$，这里

$$A_\varepsilon=\{x\in R^n|\ d(x,A)\leqslant\varepsilon\}.$$

由定理 5.23，f 在 U 上连续。因为 A_ε 是紧的，存在m，$M\in R$，使得对任何 $x\in A_\varepsilon$，$m\leqslant f(x)\leqslant M$。仿照在§5.21中定理

的证明。

23. 设 $(a, b) \in \mathbf{R}^p \times \mathbf{R}^q$. 应用 § 5.26 于函数族 $\{f_x | x \in B\}$，这里 f_x 定义成 $f_x(y) = f(x, y)$，而 B 是 $\mathbf{R}^p$ 内中心为 a 的球。

24. 设 $x \in E$. 因为 f 在 x_0 连续，我们有 $f'(x_0; x) \in \mathbf{R}$. 设 m 是 $E \times \mathbf{R}$ 中的直线

$$\{(x_0 + \lambda x, f(x_0) + \lambda f'(x_0; x)) | \lambda \in \mathbf{R}\}.$$

证明存在 $E \times \mathbf{R}$ 中包含 m 的闭超平面，真分离 epi(f) 和 m（见 § 3.9）。导出存在 $x_0' \in \partial f(x_0)$，使 $f'(x_0; x) = \langle x | x_0' \rangle$。为完成证明，利用定理5.36。

25. 利用练习24的结果。

第 六 章

1. 记

$$f^*(x') = \sup_{x \in E} \{\langle x | x' \rangle - \varphi(\|x\|)\}$$
$$= \sup_{t \geq 0} \{t \sup_{\|x\| = 1} \langle x | x' \rangle - \varphi(t)\}$$

（见 § 6.4，例 (b)）。

2. 在Fenchel不等式中取 $x = x'$（见 § 6.9）。

3. 应用定理6.15.

4. (a) 我们知道 $f^{**}(x_0) \leqslant f(x_0)$（见定理 6.11）。设 $x' \in \partial f(x_0)$。对每个 $x \in E$ 有

$$f(x) \geqslant f(x_0) + \langle x - x_0 | x_0' \rangle$$

因此

$$f^{**}(x) \geqslant f(x_0) + \langle x - x_0 | x_0' \rangle$$

（见定理6.11）．从而$f^{**}(x_0)\geqslant f(x_0)$．

（b） 利用定理6.10和（a）的结果．

5． 首先证明$f^{**}\subset\Gamma(E)$（结合定理6.11（a）和定义6.17）．

6． 应用定理6.10和6.16．

7． 若f非正常凸，则$f=+\infty$或至少存在一个x，使$f(x)=-\infty$。在这两种情形下，f^*是非正常凸的(见§6.3)．

如果f是正常凸的，则$\bar{f}$是正常凸的（见定理5.24），并且$f^{**}=\bar{f}$（见定理6.16）．导出f^*是正常凸函数．

8． $\delta_A^{**}=\mathrm{cl}(\mathrm{co}(\delta_A))=\mathrm{cl}(\delta\mathrm{co}(A))=\overline{\delta}\,\mathrm{co}(A)=\delta_B$，这里$B=\overline{\mathrm{co}(A)}$（见定理6.20）．

9． （a） 从定义6.1(a)直接得到。注意$\delta_C\Box\delta_D=\delta_{C+D}$．

（b） 利用定理6.20（e）．

10． 设f是R^n上正齐次实凸函数，f是连续的，因此$f\in\Gamma_0(R^n)$．由§6.22，存在R^n的非空闭凸子集C，使得$f=\delta_C^*$。对每个$x'\in\mathrm{R}^n$

$$\delta_C^*(x')\in R \quad 因此 \quad \sup_{x\in C}(x\mid x')<+\infty$$

导出C是有界的．

反之，设C是R^n的非空有界凸子集．δ_C^*是正齐次的和凸的。对每个$x'\in\mathrm{R}^n$及每个$x\in C$

$$(x\mid x')\leqslant M\|x'\|$$

这里$M=\sup_{x\in C}\|x\|$。从而$\delta_C^*(x')\in\mathrm{R}$．

11． 若$\bar{f}$非正常凸，则$\mathrm{cl}(f)=-\infty=\delta_\varnothing^*$。若$\bar{f}$是正常凸的，则$\mathrm{cl}(f)=\bar{f}$。应用§6.22（$\bar{f}$的正齐次性，例如，从定

理5.8（c）得出）。

12. 利用 $\partial g(0)$的定义和定理5.36。

13. 从δ_K^*和K^0的定义直接得到。

14. 存在 $a \in \mathbb{R}^n$，使 $\delta_A^*(x) = (x \mid a)$（$x \in \mathbb{R}^n$）。从而，$\delta_A^{**} = \delta_{\{a\}}$（见§6.4）。现在用练习 8 的结果。

符号汇编

$[a, b]$	作为 R 的子集：R中所有满足 $a \leqslant x \leqslant b$ 的 x 之集合； 在线性空间中：包含端点 a 和 b 的线段(见§2.1)。
$\langle a, b\rangle$	$[a, b]$的内部。$\langle a, b]$和$[a, b\rangle$有类似定义。
$(x \mid y)$	x 和 y 的内积（见§4.5）
$\langle x \mid u\rangle$	$u \in E'$ 在 $x \in E$的值（见§3.13）
$\mathbf{R}_+$	R中所有 $x \geqslant 0$ 的集合
$\overline{\mathbf{R}}$	$\mathbf{R} \cup \{+\infty\} \cup \{-\infty\}$（见§1.18）
$\mathrm{int}(A)$	A的内部
$\overline{A}$	A的闭包
$\mathrm{bd}(A)$	A的边界
$\mathrm{ri}(A)$	A的相对内部（见§4.8）
$\mathrm{rb}(A)$	A的相对边界（见§4.12）
A^i	A的代数内部（见§2.14）
A^a	A的代数闭包（见§2.14）
$\mathrm{co}(A)$	A的凸包（见§2.2）
$\overline{\mathrm{co}}(A)$	A的闭凸包（见§2.25）
$\mathrm{aff}(A)$	A的仿射包（见§2.6）
$\dim(A)$	A的维数（见§2.6）
E'	E的对偶（见§3.13）
K^0	K的极（见§3.13）

K^{00}　K的二次极（见§3.13）

P_n　$\mathbf{R}^n$的非负卦限（见§4.14）

T^t　线性映射T的伴随映射

$V\oplus\mathbf{R}$　所有（x，λ）$\in V\times\mathbf{R}$组成的线性空间（见§5.1）

$f(A)<\alpha$ 对任意$x\in A$有$f(x)<\alpha$

$\mathrm{dom}(f)$　f的有效域（见§1.21，§5.11，和§7.14）

$\mathrm{epi}(f)$　f的上图象（见§5.1）

$\bar{f}$　f的下半连续包（见§5.5）

f'_+　f的右导数

f'_-　f的左导数

$f|m$　f在m上的限制

$\mathrm{co}(f)$　f的凸包（见§5.16）

$f\square g$　f和g的下确界卷积（见§5.17）

$f'(x_0;x)$ f在x_0沿方向x的方向导数（见§5.19）

$B(a;r)$　中心是a，半径为r的闭球

∇f　f的G——微分（见§5.28）

∂f　f的次微分（见§5.30）

$\mathrm{dom}(\partial f)$ ∂f的有效域（见§5.30）

f^*　f的共轭（见§6.1）

f^{**}　f的二次极（见§6.1）

$\mathrm{cl}(f)$　f的闭包（见§6.13）

δ_A　A的指标函数（见§5.15）

δ_A^*　A的支撑函数（见§6.5）

$\Gamma(E)$　见定义6.17

$\Gamma_0(E)$　见§6.19

prox_f　关于f的邻近映射（见§7.17）

术 语 索 引

A

B

C

Carathéodory number,	Carathéodory数,
Carathéodory theorem	Carathéodory定理,
closed convex hull,	闭凸包,
closed function,	闭函数,
closed half—space.	闭半空间,
closed hyperplane,	闭超平面,
closure of a function,	函数的闭包,
concave,	凹的,
proper,	正常凹的,
cone,	锥,
convex,	凸锥,
finitely generated convex,	有限生成凸锥,
cone of supporting functionals,	支撑函数的锥,
conjugate,	共轭,
constraint qualification,	约束品性,
convex algebraic body,	凸代数体,
convex body,	凸体,
convex combination,	凸组合,
convex cone,	凸锥,
convex function,	凸函数,
improper,	非正常凸函数,
proper,	正常凸函数,
strictly,	严格凸函数,

convex hull,	凸包,
convex polytope,	凸多胞形,
convex programming,	凸规划,
convex set,	凸集:
convexity space,	凸性空间,
cyclically monotone,	循环单调,
maximal,	最大循环单调,

D

derivative,	导数,
differentiable,	可微的,
Fréchet—,	Fréchet—可微,
Gateaux—,	Gateaux—可微,
dimension,	维数,
directional derivative,	方向导数,
domain,	定义域,
effective,	有效域,
dual of a function,	函数的对偶,
dual of a normed linear space,	赋范线性空间的对偶,
dual problem,	对偶问题,
duality theorem of Fenchel,	Fenchel对偶定理,

E

effective domain,	有效域,
epigraph,	上图象,

equality constraints, 等式约束，
extreme point, 极端点，

F

Farkas' lemma, Farkas引理，
feasible solution, 可行解，
  strictly, 严格可行解，
Fenchel's duality theorem, Fenchel对偶定理，
Fenchel's inequality, Fenchel不等式。
finitely generated convex cone, 有限生成凸锥，
Fréchet derivative, Fréchet导数，
Fréchet-differentiable, Fréchet-可微，

G

Gateaux-differentiable, Gateaux-可微，
Gateaux-differential, Gateaux-微分，
gauge, 度规，
generator, 生成元，
Gordan's lemma, Gordan引理，

H

Hahn-Banach theorem, Hahn-Banach定理，
Hausdorff distance, Hausdorff距离，

Helly number,	Helly数，
Helly's theorem.	Helly定理，
Hölder's inequality,	Hölder不等式，
hyperplane,	超平面，
closed,	闭超平面.
non-vertical,	非垂直闭超平面，
supporting,	支撑超平面，
vertical,	垂直超平面，

I

improper convex,	非正常凸，
indicator function,	指标函数.
inequality constraints,	不等式约束，
inequality of Fenchel,	Fenchel不等式，
inequality of Hölder,	Hölder不等式，
inequality of Jensen,	Jensen不等式，
inequality of Young,	Young不等式，
infimal convolution,	下确界卷积，
interior of a line segment.	线段的内部，

J

Jensen's inequality,	Jensen不等式，

K

k-simplex, *k*-单纯形,
Kirchberger's theorem, Kirchberger定理,
Kuhn-Tucker conditons, Kuhn-Tucker条件,

L

Lagrange function, 拉格朗日函数,
Lagrange multipliers, 拉格朗日乘子,
Lagrangian, 拉格朗日算子,
line segment, 线段,
linear programming, 线性规划,
linear topological space, 线性拓扑空间,
Lipschitzian, 李普西兹,
  locally, 局部李普西兹,
  locally equi- 局部等度李普西兹,
locally bounded, 局部有界的,
locally convex space, 局部凸空间,
logarithmically convex, 对数凸的,
lower semi-continuity, 下半连续性,
lower semi-continuous hull, 下半连续包,

M

maximal (cyclically) monotone, 最大(循环)单调,

midpoint convex,	中点凸的,
Minkowski, theorem of,	Minkowski定理,
Minkowski distance functional,	Minkowski距离函数,
minorant,	弱的,
monotone operator,	单调算子,
cyclically,	循环单调算子,
maximal (cyclically),	最大(循环)单调算子,
multifunction,	多值函数,
convex,	凸多值函数,
multipliers,	乘子,

N

non-trivial supporting hyperplane,	非平凡支撑超平面,
non-vertical hyperplane,	非垂直超平面,
norm topology,	范拓扑,
normal,	正规的,
normal cone,	正规锥,
normed linear space,	正规线性空间,

O

optimal solution,	最优解,

P

polar of a cone,	锥的极,

polar of a function,	函数的极，
polyhedral cone,	多面锥，
positively homogeneous,	正齐次的，
projection,	投影，
proper concave,	正常凹的，
proper convex,	正常凸的，
proper separation,	真分离，
proximity mapping,	邻近映射，

Q

quasi-convex,	拟凸的，
strictly,	严格拟凸的，

R

Radon number,	Radon数，
Radon's theorem,	Radon定理，
relative boundary,	相对边界，
relative interior,	相对内部，

S

saddle point,	鞍点，
separation,	分离，
proper,	真分离，
strict,	严格分离，

separation theorem,	分离定理,
simplex,	单纯形,
Slater's condition,	Slater条件,
star-shaped,	星形的,
strict separation,	严格分离,
strictly convex,	严格凸,
strictly feasible solution,	严格可行解,
strictly quasi-convex,	严格拟凸的,
subadditive,	次加性的,
subdifferentiable,	次可微的,
subdifferential,	次微分,
subgradient,	次梯度,
supporting function,	支撑函数,
supporting hyperplane,	支撑超平面,
non-trivial,	非平凡的支撑超平面,

T

theorem of the alternative,	择一定理,

V

vertex,	顶点,
vertical hyperplane,	垂直超平面,

W

weak topology, 弱拓扑,

Y

Young's inequality, Young不等式,